









Qy 18 ArgAlaaspSe-SerGluGlulysArgHisArgLysargLysLysHisHisargGlyTyr 37  
 nb 3779614 CCGGCCGCTTCGTCGCCGTACCGCCGCGCACCGCGCTTCGCGCTACCGCCGACGGAAC 3779955

## Alignment Scores:

Pred. No.: 9.72 Length: 750  
 Score: 63.00 Matches: 13  
 Percent Similarity: 58.33% Conservative: 1  
 Best Local Similarity: 54.17% Mismatches: 10  
 Query Match: 19.94% Indels: 0  
 DB: 4 Gaps: 0

US-10-079-754A-10 (1-58) x US-09-359-301A-30 (1-750)

QY 35 ArgGlyTyrPheGlnGlnTyrGlnProTyrGlnArgTyrProLeuAsnTyrProProAla 54

DB 199 CAGGATATCTCCACACAGGTATCTCTCAGCAAGATATCTTCACCGTAGCGGCTCAA 258

QY 55 TyrProPhePro 58

DB 259 TATCTCTCCACCA 270

## RESULT 13

US-08-308-893-1

; Sequence 1, Application US/08308883

; Patent No. 5576300

; GENERAL INFORMATION:

; APPLICANT: Mukerji, P.

; APPLICANT: Prieto, P. A.

; APPLICANT: Seo, A. E.-Y.

; APPLICANT: Baxter, J. H.

; APPLICANT: Cummings, R.D.

; TITLE OF INVENTION: Method for Inhibition of Human Rotavirus Infection.

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lonnie R. Drayer

; ADDRESSEE: ROSS Products Division

; ADDRESSEE: Abbott Laboratories

; STREET: 625 Cleveland Avenue

; CITY: Columbus

; STATE: Ohio

; COUNTRY: United States

; ZIP: 43215

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb storage

; COMPUTER: Apple Macintosh

; OPERATING SYSTEM: Macintosh System 7.1

; SOFTWARE: ClarisWorks 1.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/308,883

; FILING DATE: 16-SEP-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA: No. 5576300 applicable

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (614) 624-3774

; TELEFAX: (614) 624-3074

; TELEX: No. 5576300e

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 857 base pairs

; TYPE: Nucleic acid

; STRANDEDNESS: Single

; TOPOLOGY: Linear

; MOLECULE TYPE: cDNA

; DESCRIPTION: Human milk kappa-casein

; HYPOTHETICAL: No

; ANTI-SENSE:

; FRAGMENT TYPE:

; ORIGINAL SOURCE: Human

; ORGANISM: Homo sapiens

; STRAIN:

; INDIVIDUAL ISOLATE:

; DEVELOPMENTAL STAGE: Adult

; HAPLOTYPE:

; TISSUE TYPE: Mammary gland

; CELL TYPE:

; CELL LINE:

; ORGANELLE:

; IMMEDIATE SOURCE: Human Mammary Gland

; LIBRARY:

; CLONE:

; POSITION IN GENOME:

; CHROMOSOME/SEGMENT:

; MAP POSITION:

; UNITS:

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 45...593

; IDENTIFICATION METHOD: DNA sequencing and restriction analysis

; OTHER INFORMATION: The encoded product of nucleotide SEQ ID NO: 1: is the human

; PUBLICATION INFORMATION:

; AUTHORS: L. Hanson et al

; TITLE: DNA Encoding Kappa-Casein, Process for Obtaining the Protein and Use Ther

; JOURNAL:

; VOLUME:

; ISSUE:

; PAGES:

; DATE:

; DOCUMENT NUMBER: PCT/WO93/15196

; FILING DATE: 25-JAN-1993

; PUBLICATION DATE: 05-AUG-1993

; RELEVANT RESIDUES IN SEQ ID NO:

; US-08-308-883-1

; Alignment Scores:

; Pred. No.: 11.7 Length: 857

; Score: 63.00 Matches: 17

; Percent Similarity: 50.85% Conservative: 13

; Best Local Similarity: 28.81% Mismatches: 25

; Query Match: 19.94% Indels: 4

; DB: 1 Gaps: 3

; US-10-079-754A-10 (1-58) x US-08-308-893-1 (1-857)

QY 1 MetLysIlePheHellePheValPhe---IleMetAlaLeuIleLeuAlaMetIleArgAla 19

DB 45 ATGAAGAGT---TCTCTAGTTGTCATGCGCTGGCATTACCTGCTGCTTTTGGCTGTG 104

QY 20 AspSerSerGluGluLysArgHisArgLysArgLysHisArgGlyTyrPheGln 39

DB 105 GAGGTTCAAACCCAGAAACACACAGCAGTCGCCATGAGAAATGATGAAGACCACTTCTACAG 164

QY 40 GlnTyrGlnProTyrGlnArgTyrProLeuAsnTyr---ProProAlaTyrProPhe 57

DB 165 AAACACAGCTCCATAT-----GTCCCAATGTATTATGTGCCAAATAGTATCCTTAT 215

RESULT 14

US-08-730-163-1

; Sequence 1, Application US/08730163

; Patent No. 5712250

; GENERAL INFORMATION:

; APPLICANT: Mukerji, P.

; APPLICANT: Prieto, P. A.

; APPLICANT: Seo, A. E.-Y.

; APPLICANT: Baxter, J. H.

; APPLICANT: Cummings, R.D.

; TITLE OF INVENTION: Product for Inhibition of Human Rotavirus Infection.

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lonnie R. Drayer

; ADDRESSEE: ROSS Products Division

; ADDRESSEE: Abbott Laboratories

; STREET: 625 Cleveland Avenue

; CITY: Columbus

; STATE: Ohio

; COUNTRY: United States

; ZIP: 43215

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb storage (B) COMPUTER: Apple Macintosh

; OPERATING SYSTEM: Macintosh System 7.1 (D) SOFTWARE: ClarisWorks 1.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/730,163  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/308,882  
FILING DATE: 16-SEP-1994  
TELEPHONE: (614) 624-3774  
TELEFAX: (614) 624-3074  
TELEX: NO. 5712250e  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 857 base pairs  
TYPE: Nucleic acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: CDNA  
DESCRIPTION: Human milk kappa-casein  
HYPOTHETICAL: No  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE: Human  
ORGANISM: Homo sapiens  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE: Adult  
HAPLOTYPE:  
TISSUE TYPE: Mammary gland  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE: Human Mammary Gland  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 45...593  
IDENTIFICATION METHOD: DNA sequencing and restriction analysis  
OTHER INFORMATION: The encoded product of nucleotide SEQ ID NO: 1: is the human  
PUBLICATION INFORMATION:  
AUTHORS: L. Hansson et al  
TITLE: DNA Encoding Kappa-Casein, Process for Obtaining the Protein and Use The  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER: PCT/WO93/15196  
FILING DATE: 25-JAN-1993  
PUBLICATION DATE: 05-AUG-1993  
RELEVANT RESIDUES IN SEQ ID NO:  
US-08-730-163-1

Alignment Scores:  
Pred. No.: 11.7 Length: 857  
Score: 63.00 Matches: 17  
Percent Similarity: 50.85% Conservative: 13  
Best Local Similarity: 28.81% Mismatches: 25  
Query Match: 19.94% Indels: 4  
DB: 1 Gaps: 3

US-10-079-754A-10 (1-58) x US-08-730-163-1 (1-857)

Qy 1 MetLysIlePheIlePheValPhe---IleMetAlaLeuIleLeuAlaMetIleArgAla 19  
Db 45 ARGAGAGTTTCTCTAGTGTCAATGCCCTGGCAITTAACCCGCTTTTTCGTGTG 104  
Qy 20 AspSerSerGluGluLysArgHisArgLysArgLysLysHisArgGlyTyrPheGln 39

Db 105 GAGGTTCACCAACCAACACAGCAGTCATGAGATGATGAAGACCATTTCTATCAG 164  
Qy 40 GlnTyrGlnProTyrGlnArgTyrProLeuAsnTyr---ProPheAlaTyrProPhe 57  
Db 165 AAACAGCTCCATAT-----GTCCCAATGATTATTATGTGCCAAATAGCTATCCTTAT 215  
RESULT 15  
US-08-256-799-1  
Sequence 1, Application US/08256799  
Patent No. 6222094  
GENERAL INFORMATION:  
APPLICANT: HANSSON, Lennart  
APPLICANT: STROEMQVIST, Mats  
APPLICANT: BERGSTROEM, Sven  
APPLICANT: HERNELL, Olle  
APPLICANT: Toernell, Jan  
TITLE OF INVENTION: DNA ENCODING KAPPA-CASEIN, PROCESS FOR  
TITLE OF INVENTION: OBTAINING THE PROTEIN AND USE THEREOF  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSER: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,799  
FILING DATE: 06-DEC-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DK 88/92  
FILING DATE: 23-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, Iver P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: HANSSON=1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 857 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 45...593  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 45...593  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 45...104  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 13..44  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 594...848  
US-08-256-799-1

Alignment Scores:  
Pred. No.: 11.7 Length: 857  
Score: 63.00 Matches: 17  
Percent Similarity: 50.85% Conservative: 13  
Best Local Similarity: 28.81% Mismatches: 25  
Query Match: 19.94% Indels: 4  
DB: 3 Gaps: 3

US-10-079-754A-10 (1-58) x US-08-256-799-1 (1-857)  
Qy 1 MetLysIlePheIlePheValPhe--IleMetAlaLeuIleLeuAlaMetIleArgAla 19  
Db 45 ATGAAGAGTTTCTCTAGTTGTCAATGCCCTGCCATTAAACCTTCCTTTTGGCTGTG 104  
Qy 20 AspSerSerGluGluLysArgHisArgLysArgLysLysHisArgGlyTyrPheGln 39  
Db 105 GAGGTTCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 164  
Qy 40 GlnTyrGlnProTyrGlnArgTyrProLeuAsnTyr---ProProAlaTyrProPhe 57  
Db 165 AAAACAGCTCCATAT-----GTCCCAATGATTATTGTGCCAAATAGCTATCCTTAT 215

Search completed: August 16, 2004, 00:49:11  
Job time : 1484 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame plus p2n model

Run on: August 16, 2004, 00:23:24 ; Search time 326 Seconds  
(without alignments) 872.955 Million cell updates/sec

**Title:** US-10-079-754A-10

Perfect score: 316  
Sequence: 1 MKIFIFVFMALILAMIRAD.....QQYQYQYVPLNYPDPVFPF 58

Scoring table: BLOSUM62

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

```
Minimum DB seq length: 0
```

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

## Listing first 45 summaries

Command line parameters:

```

MODEL=frame_p2n_model_devs-xlh
-O=/cgn2.1/USPTO.spool/US10079754/runat 06082004 181230 29453/app_query.fasta.1199
-DB=Published Applications NA -GPMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCLD=0 -LOOPPEXT=0 -UNITS=bits -SPART=1 -END=-1 -MATRIX=blousem2
-TRANS=human40 cdi -LIST=45 -DOCALIGN=200 -THR SCORE=opt -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODS=LOCAL -OUTFMF=ptc -NOSM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 USER=US10079754_CGN 1 480 @runat 06082004 181230 29453
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY=NO SCORES=0 -WAIT -DPSLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-EGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database : Published Applications NA: \*\*

- 1: /cgn2\_6/pdata/2/pubppna/us07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/pdata/2/pubppna/pct\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/pdata/2/pubppna/us06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/pdata/2/pubppna/us06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/pdata/2/pubppna/us07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/pdata/2/pubppna/pctus\_PUBCOMB.seq.\*
- 7: /cgn2\_6/pdata/2/pubppna/us08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/pdata/2/pubppna/us08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/pdata/2/pubppna/us09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/pdata/2/pubppna/us09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/pdata/2/pubppna/us09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/pdata/2/pubppna/us09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/pdata/2/pubppna/us09\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/pdata/2/pubppna/us10A\_PUBCOMB.seq.\*
- 15: /cgn2\_6/pdata/2/pubppna/us10B\_PUBCOMB.seq.\*
- 16: /cgn2\_6/pdata/2/pubppna/us10C\_PUBCOMB.seq.\*
- 17: /cgn2\_6/pdata/2/pubppna/us10C\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/pdata/2/pubppna/us60\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/pdata/2/pubppna/us60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	100	100	100	100	100	100

## ALIGNMENTS

RESULT 1

```

US-10-079-754A-15
? Sequence 15, Application US/10079754A
? Publication No. US20020164625A1
? GENERAL INFORMATION:
? APPLICANT: Glenn, Matthew
? APPLICANT: Grigor, Murray R.
? APPLICANT: Molenaar, Adrian J.
? APPLICANT: Davis, Stephen R.
? TITLE OF INVENTION: Compositions Isolated from Bovine
? TITLE OF INVENTION: Mammary Gland and Methods for Their Use
? FILE REFERENCE: 1100C.1068
? CURRENT APPLICATION NUMBER: US/10/079,754A
? CURRENT FILING DATE: 2002-02-19
? PRIOR APPLICATION NUMBER: US 09/599,146
? PRIOR FILING DATE: 2000-10-27
? PRIOR APPLICATION NUMBER: US 60,162,701
? PRIOR FILING DATE: 1999-10-29
? PRIOR APPLICATION NUMBER: US 09/644,190
? PRIOR FILING DATE: 2000-08-22
? PRIOR APPLICATION NUMBER: US 60,150,330
? PRIOR FILING DATE: 1999-08-23
? NUMBER OF SEQ ID NOS: 15

```

```

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Bovine
US-10-079-754A-15
Alignment Scores:
Pred. No.: 7,07e-37 Length: 267
Score: 316.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-079-754A-10 (1-58) x US-10-079-754A-15 (1-267)
QY 1 MetLysIlePheIlePheValPheIleMetAlaLeuIleLeuAlaMetIleArgAlaAsp 20
DB 44 ATGAAGATCTTTATCTTTGCTTCATTATGCTCTCATCTAGCCATGATTAGAGCTGAT 103
QY 21 SerSerGluGluLysArgHisArgLysArgLysLysHisArgGlyTyrPheGlnGln 40
DB 104 TCATCTGAAGAGAAACGTCACAGGAACCGAAAAACATCATAGAGGATATTTCAACAA 163
QY 41 TyrGlnProTyrGlnArgTyrProLeuAsnTyrProAlaTyrProPhePro 58
DB 164 TACCAGCCATATCAACGATATCCAAATAATATCTCTCGGTATCCATTTCCT 217

RESULT 2
US-10-079-623-200
; Sequence 200, Application US/10079623
; Publication No. US20020169302A1
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; TITLE OF INVENTION: Compositions isolated from bovine
; FILE REFERENCE: 11000.1044C3
; CURRENT APPLICATION NUMBER: US/10/079,623
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 200
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Bovine
US-10-079-623-200
Alignment Scores:
Pred. No.: 7,07e-37 Length: 267
Score: 316.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-079-754A-10 (1-58) x US-10-079-623-200 (1-267)
QY 1 MetLysIlePheIlePheValPheIleMetAlaLeuIleLeuAlaMetIleArgAlaAsp 20
DB 44 ATGAAGATCTTTATCTTTGCTTCATTATGCTCTCATCTAGCCATGATTAGAGCTGAT 103
QY 21 SerSerGluGluLysArgHisArgLysArgLysLysHisArgGlyTyrPheGlnGln 40
DB 104 TCATCTGAAGAGAAACGTCACAGGAACCGAAAAACATCATAGAGGATATTTCAACAA 163
QY 41 TyrGlnProTyrGlnArgTyrProLeuAsnTyrProAlaTyrProPhePro 58
DB 164 TACCAGCCATATCAACGATATCCAAATAATATCTCTCGGTATCCATTTCCT 217

RESULT 3
US-10-079-754A-1
; Sequence 1, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions isolated from Bovine
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Bovine
US-10-079-754A-1
Alignment Scores:
Pred. No.: 1.68e-36 Length: 505
Score: 316.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-079-754A-10 (1-58) x US-10-079-754A-1 (1-505)
QY 1 MetLysIlePheIlePheValPheIleMetAlaLeuIleLeuAlaMetIleArgAlaAsp 20
DB 44 ATGAAGATCTTTATCTTTGCTTCATTATGCTCTCATCTAGCCATGATTAGAGCTGAT 103
QY 21 SerSerGluGluLysArgHisArgLysArgLysLysHisArgGlyTyrPheGlnGln 40
DB 104 TCATCTGAAGAGAAACGTCACAGGAACCGAAAAACATCATAGAGGATATTTCAACAA 163
QY 41 TyrGlnProTyrGlnArgTyrProLeuAsnTyrProAlaTyrProPhePro 58
DB 164 TACCAGCCATATCAACGATATCCAAATAATATCTCTCGGTATCCATTTCCT 217

RESULT 4
US-10-079-754A-4
; Sequence 4, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions isolated from Bovine
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
```

US-10-079-754A-10 (1-58) x US-10-079-754A-2 (1-585)  
Qy 1 MetLysilePheIlePheValPheIleVerAlalaLeuleuAlaMetIleArgAlaAsp 20  
| | | | | . . . . .  
; Publication No. US20020164625A1  
; GENERAL INFORMATION:  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Grigor, Murray R.



;; PRIOR APPLICATION NUMBER: 60/311,979  
;; PRIOR FILING DATE: 2001-08-13  
;; PRIOR APPLICATION NUMBER: 60/313,201  
;; PRIOR FILING DATE: 2001-08-17  
;; PRIOR APPLICATION NUMBER: 60/312,892  
;; PRIOR FILING DATE: 2001-08-16  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 327  
;; SOFTWARE: Curasequlst version 0.1  
;; SEQ ID NO 167  
;; LENGTH: 678  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (270)..(455)  
US-10-210-172-167  
Alignment Scores:  
Pred. No.: 2,79e-06 Length: 678  
Score: 110.00 Matches: 27  
Percent Similarity: 61.82% Conservative: 7  
Best Local Similarity: 49.09% Mismatches: 19  
Query Match: 34.81% Indels: 2  
DB: 13 Gaps: 2  
US-10-079-754A-10 (1-58) x US-10-210-172-167 (1-678)  
Qy 1 MetLysIlePheIlePheValPheIleMetAlaLeuIleLeuAlaMetIleArgAlaAsp 20  
Db 270 ATGAAGTTCCTTGGCTTCATCTTCGCTCTCATGTTCCATGATTCGAGCTGAT 329  
Qy 21 SerSerGluGluLysArgHisArgLysArgLysLysHisArgGlyTyrPheGlnGln 40  
Db 330 TCATCTCGAAGAAATTTTGGGTAGAAATTCGAGATTCGGTTATGGGTAT---GGCCCT 386  
Qy 41 TyrGlnProTyrGlnArgTyrProLeuAsnTyrProProAlaTyr 55  
Db 387 TATCAGCCAGTTCAGAAACACCACTA---TACCCACACCATAC 428  
RESULT 10  
US-10-104-047-1193  
;; Sequence 1193, Application US/10104047  
;; Publication No. US20030236392A1  
;; GENERAL INFORMATION:  
;; APPLICANT: HELIX RESEARCH INSTITUTE  
;; TITLE OF INVENTION: NO. US20030236392A1el full length cDNA  
;; FILE REFERENCE: H1-A0105  
;; CURRENT APPLICATION NUMBER: US/10/104,047  
;; CURRENT FILING DATE: 2002-03-25  
;; PRIOR APPLICATION NUMBER:  
;; PRIOR FILING DATE:  
;; NUMBER OF SEQ ID NOS: 4096  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 1193  
;; LENGTH: 1584  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-104-047-1193  
Alignment Scores:  
Pred. No.: 8.88e-06 Length: 1584  
Score: 110.00 Matches: 27  
Percent Similarity: 61.82% Conservative: 7  
Best Local Similarity: 49.09% Mismatches: 19  
Query Match: 34.81% Indels: 2  
DB: 16 Gaps: 2  
US-10-079-754A-10 (1-58) x US-10-104-047-1193 (1-1584)  
Qy 1 MetLysIlePheIlePheValPheIleMetAlaLeuIleLeuAlaMetIleArgAlaAsp 20  
Db 1097 ATGAAGTTCCTTGGCTTCATCTTCGCTCTCATGTTCCATGATTCGAGCTGAT 1156

Qy 21 SerSerGluGluLysArgHisArgLysArgLysLysHisArgGlyTyrPheGlnGln 40  
Db 1157 TCATCTCGAAGAAATTTTGGGTAGAAATTCGAGATTCGGTTATGGGTAT---GGCCCT 1213  
Qy 41 TyrGlnProTyrGlnArgTyrProLeuAsnTyrProProAlaTyr 55  
Db 1214 TATCAGCCAGTTCAGAAACACCACTA---TACCCACACCATAC 1255  
RESULT 11  
US-09-992-600A-5  
;; Sequence 5, Application US/09992600A  
;; Publication No. US20030027161A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Benjanin, Stephanie  
;; APPLICANT: Tanaka, Hiroaki  
;; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
;; FILE REFERENCE: 91 US4 DIV  
;; CURRENT APPLICATION NUMBER: US/09/992,600A  
;; CURRENT FILING DATE: 2001-11-13  
;; PRIOR APPLICATION NUMBER: US 09/924,340  
;; PRIOR FILING DATE: 2001-08-06  
;; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
;; PRIOR FILING DATE: 2001-08-06  
;; PRIOR APPLICATION NUMBER: US 60/305,456  
;; PRIOR FILING DATE: 2001-07-13  
;; PRIOR APPLICATION NUMBER: US 60/302,277  
;; PRIOR FILING DATE: 2001-06-29  
;; PRIOR APPLICATION NUMBER: US 60/298,698  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: US 60/293,574  
;; PRIOR FILING DATE: 2001-05-25  
;; NUMBER OF SEQ ID NOS: 114  
;; SOFTWARE: Jpatent  
;; SEQ ID NO 5  
;; LENGTH: 438  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: 5'UTR  
;; LOCATION: 1..83  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 84..317  
;; FEATURE:  
;; NAME/KEY: 3'UTR  
;; LOCATION: 318..438  
;; FEATURE:  
;; NAME/KEY: polyA signal  
;; LOCATION: 397..402  
;; FEATURE:  
;; NAME/KEY: polyA site  
;; LOCATION: 423..438  
US-09-992-600A-5  
Alignment Scores:  
Pred. No.: 0.000281 Length: 438  
Score: 94.50 Matches: 18  
Percent Similarity: 56.36% Conservative: 13  
Best Local Similarity: 32.73% Mismatches: 17  
Query Match: 29.91% Indels: 7  
DB: 10 Gaps: 1  
US-10-079-754A-10 (1-58) x US-09-992-600A-5 (1-438)  
Qy 1 MetLysIlePheIlePheValPheIleMetAlaLeuIleLeuAlaMetIleArgAlaAsp 20  
Db 84 ATGAAGTTCCTTGGCTTCATCTTCGCTCTCATGTTCCATGATTCGAGCTGAT 143  
Qy 21 SerSerGluGluLysArgHisArgLysArgLysLysHisArgGlyTyrPheGlnGln 40  
Db 144 TCATCGAAAGAGACATCATCGGTATAGAGAAATTCAT----- 185









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2004, 17:06:10 ; Search time 18 Seconds  
(without alignments)  
166.350 Million cell updates/sec

Title: US-10-079-754A-10

Perfect score: 58

Sequence: 1 MKIFIFVFMALILMIRAD.....QQYQYQYPLNYPYPAYPPF 58

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PTUS.COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	10.3	12	3	US-08-602-999A-268
2	6	10.3	12	4	US-09-500-124-268
3	6	10.3	95	4	US-09-621-976-7261
4	6	10.3	103	4	US-09-460-384-35
5	6	10.3	107	4	US-09-198-452A-1200
6	6	10.3	134	4	US-09-634-238-376
7	6	10.3	146	4	US-09-489-039A-7386
8	6	10.3	164	4	US-09-134-000C-3551
9	6	10.3	201	4	US-09-543-681A-4214
10	6	10.3	216	4	US-09-543-681A-5885
11	6	10.3	241	4	US-09-489-039A-8060
12	6	10.3	263	4	US-09-134-001C-3062
13	6	10.3	298	4	US-09-543-681A-5642
14	6	10.3	310	3	US-08-605-284B-16
15	6	10.3	380	3	US-09-097-889-25
16	6	10.3	386	4	US-09-098-079-25
17	6	10.3	386	4	US-09-489-039A-8756
18	6	10.3	450	4	US-09-369-247-97
19	6	10.3	508	4	US-09-369-247-167
20	6	10.3	511	4	US-09-543-681A-6490
21	6	10.3	511	4	US-09-134-000C-5362
22	6	10.3	539	4	US-09-291-922-26
23	6	10.3	581	4	US-09-489-039A-10559
24	6	10.3	590	4	US-09-134-001C-4390
25	6	10.3	601	4	US-09-336-643A-4
26	6	10.3	623	4	US-09-540-236-1934
27	6	10.3	751	4	US-09-252-991A-29893

28	6	10.3	1026	1	US-07-998-003A-95	Sequence 95, Appl
29	6	10.3	1026	1	US-08-453-274B-95	Sequence 95, Appl
30	6	10.3	1026	1	US-08-453-695A-95	Sequence 95, Appl
31	6	10.3	1026	1	US-08-268-161A-95	Sequence 95, Appl
32	6	10.3	1026	2	US-08-453-702A-95	Sequence 95, Appl
33	6	10.3	1026	3	US-09-099-639-95	Sequence 95, Appl
34	6	10.3	1026	5	PCT-US93-12588-95	Sequence 95, Appl
35	6	10.3	1026	5	PCT-US95-08071-95	Sequence 95, Appl
36	6	10.3	1203	1	US-07-998-003A-103	Sequence 103, App
37	6	10.3	1203	1	US-08-453-274B-103	Sequence 103, App
38	6	10.3	1203	1	US-08-453-695A-103	Sequence 103, App
39	6	10.3	1203	1	US-08-268-161A-103	Sequence 103, App
40	6	10.3	1203	2	US-08-453-702A-103	Sequence 103, App
41	6	10.3	1203	3	US-09-099-639-103	Sequence 103, App
42	6	10.3	1203	5	PCT-US93-12588-103	Sequence 103, App
43	6	10.3	1203	5	PCT-US95-08071-103	Sequence 103, App
44	6	10.3	1349	3	US-08-938-291A-6	Sequence 6, Appli
45	6	10.3	1349	4	US-09-589-619-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1  
US-08-602-999A-268  
; Sequence 268, Application US/08602999A  
; Patent No. 6184205  
; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.  
; APPLICANT: KAY, Brian K.  
; APPLICANT: THORN, Judith M.  
; APPLICANT: QUILLIAM, Lawrence A.  
; APPLICANT: DER, Channing J.  
; APPLICANT: FOWLES, Dana M.  
; APPLICANT: RIDER, James E.  
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF ISOLATING AND USING SAME  
; NUMBER OF SEQUENCES: 467  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/602,999A  
; FILING DATE: 16-FEB-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Misrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-202  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 750-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 268:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-08-602-999A-268

Query Match 10.3%; Score 6; DB 3; Length 12;  
Best Local Similarity 100.0%; Pred.No. 6.6;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 YPPAYP 56  
|||||  
Db 1 YPPAYP 6

RESULT 2  
US-09-500-124-268  
; Sequence 268, Application US/09500124  
; Patent No. 6432920  
; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.  
; APPLICANT: KAY, Brian K.  
; APPLICANT: THORN, Judith M.  
; APPLICANT: CULLIAM, Lawrence A.  
; APPLICANT: DER, Channing J.  
; APPLICANT: FOWLKES, Dana M.  
; APPLICANT: RIDER, James E.  
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
; TITLE OF INVENTION: ISOLATING AND USING SAME  
; NUMBER OF SEQUENCES: 467  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/500,124  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/602,999  
; FILING DATE: 16-FEB-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-202  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 268:  
; SEQUENCE CHARACTERISTICS:  
; TYPE: amino acids  
; LENGTH: 12 amino acids  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-09-500-124-268

Query Match 10.3%; Score 6; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 YPPAYP 56  
|||||  
Db 1 YPPAYP 6

RESULT 3  
US-09-621-976-7261  
; Sequence 7261, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 7261  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-621-976-7261

Query Match 10.3%; Score 6; DB 4; Length 95;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 SEEKRH 27  
|||||  
Db 86 SEEKRH 91

## RESULT 4

US-09-460-384-35  
; Sequence 35, Application US/09460384  
; Patent No. 6337316

## ; GENERAL INFORMATION:

; APPLICANT: EL TAYAR, Nabil  
; BLECHNER, Steven  
; JAMESON, Brad  
; TEPPER, Mark

; TITLE OF INVENTION: CD28/CTLA-4 INHIBITING PEPTIDOMIMETICS,  
; PHARMACEUTICAL COMPOSITIONS THEREOF, AND METHOD OF USING  
; SAME

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.

; STREET: 624 Ninth Street N.W., Ste. 300

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20001

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/460,384

; FILING DATE: 13-Dec-1999

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US98/12312

; FILING DATE: 11-JUN-1998

; APPLICATION NUMBER: US 60/049,470

; FILING DATE: 12-JUN-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: YUN, Allen C

; REGISTRATION NUMBER: 37,971

; REFERENCE/DOCKET NUMBER: EL TAYAR=1A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 628-5197

; TELEFAX: (202) 737-3528

; INFORMATION FOR SEQ ID NO: 35:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 103 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 35:

US-09-460-384-35

Query Match 10.3%; Score 6; DB 4; Length 103;

```

Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 QQYQPY 44
Db 88 QQYQPY 93

RESULT 5
US-09-198-452A-1200
; Sequence 1200, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1200
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1200

Query Match 10.3%; Score 6; DB 4; Length 107;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 SSEEKR 26
Db 80 SSEEKR 85

RESULT 6
US-09-634-238-376
; Sequence 376, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Iixka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
; APPLICANT: Christensson, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; TITLE OF INVENTION: them and methods for using them.
; FILE REFERENCE: 11000.1043UI
; CURRENT APPLICATION NUMBER: US/09/634,238
; CURRENT FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 376
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-376

Query Match 10.3%; Score 6; DB 4; Length 134;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 YPPAYP 56
Db 52 YPPAYP 57

```

```

RESULT 7
US-09-489-039A-7386
; Sequence 7386, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7386
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7386

Query Match 10.3%; Score 6; DB 4; Length 146;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 RADSSS 23
Db 90 RADSSS 95

RESULT 8
US-09-134-000C-3551
; Sequence 3551, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3551
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3551

Query Match 10.3%; Score 6; DB 4; Length 164;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 SSEEKR 26
Db 151 SSEEKR 156

RESULT 9
US-09-543-681A-4214
; Sequence 4214, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344

```

```
; SEQ ID NO 4214
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4214

Query Match      10.3%; Score 6; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      10 MALILA 15
Db      105 MALILA 110

RESULT 10
US-09-543-681A-5885
; Sequence 5885, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5885
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5885

Query Match      10.3%; Score 6; DB 4; Length 216;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 FIMALI 13
Db      28 FIMALI 33

RESULT 11
US-09-489-039A-8060
; Sequence 8060, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8060
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8060

Query Match      10.3%; Score 6; DB 4; Length 241;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 FIMALI 13
Db      183 FIMALI 188

; SEQ ID NO 4214
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4214

Query Match      10.3%; Score 6; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      10 MALILA 15
Db      105 MALILA 110

RESULT 12
US-09-134-001C-3062
; Sequence 3062, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134.001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3062
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3062

Query Match      10.3%; Score 6; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 FIFVEI 9
Db      10 FIFVEI 15

RESULT 13
US-09-543-681A-5642
; Sequence 5642, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5642
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5642

Query Match      10.3%; Score 6; DB 4; Length 298;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      21 SSEKR 26
Db      264 SSEKR 269

RESULT 14
US-08-605-284B-16
; Sequence 16, Application US/08605284B
; Patent No. 6050271
; GENERAL INFORMATION:
; APPLICANT: WALEWSKI, JOSE L.
; APPLICANT: RECIO-PINTO, ESPERANZA
; TITLE OF INVENTION: VOLTAGE GATED SODIUM CHANNELS FROM
; TITLE OF INVENTION: HUMAN PERIPHERAL NERVE
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
; STREET: CLINTON SQUARE, P.O. BOX 1051
```

```

; CITY: ROCHESTER
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,284B
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: BRAMAN, SUSAN J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/800 (CRF D-1705)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1636
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-605-284B-16

Query Match      10.3%; Score 6; DB 3; Length 310;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 FINALI 13
      |||||
Db      67 FINALI 72

RESULT 15
US-09-097-889-25
; Sequence 25, Application US/09097889
; Patent No. 6218117
; GENERAL INFORMATION:
; APPLICANT: Heintzstadt, Corrina
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: AGENTS THAT QUANTATIVELY ALTER DETECTABLE
; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA: MITOCHONDRIAL DNA RATIOS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,889
; FILING DATE: 15-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rosenman Ph.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 660088.417
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-09-097-889-25

Query Match      10.3%; Score 6; DB 3; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      12 LILAMI 17
      |||||
Db      299 LILAMI 304

Search completed: August 6, 2004, 17:10:28
Job time : 23 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2004, 17:09:05 ; Search time 46 Seconds  
(without alignments)  
395.513 Million cell updates/sec

Title: US-10-079-754A-10

Perfect score: 58

Sequence: 1 MKKIFVFMALILAMIRAD.....QQYQYQRYPLNYPAYPP 58

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 1291235 seqs, 313682936 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	100.0	58	13	US-10-079-754A-7
2	58	100.0	58	13	US-10-079-754A-10
3	51	87.9	58	13	US-10-079-754A-8
4	34	58.6	59	13	US-10-079-754A-12
5	27	46.6	70	13	US-10-079-754A-9
6	11	19.0	21	13	US-10-079-754A-11
7	8	13.8	363	12	US-10-425-114-70822
8	7	12.1	13	10	US-09-876-904A-586
9	7	12.1	62	12	US-10-210-172-168
10	7	12.1	72	12	US-10-424-599-145633
11	7	12.1	150	12	US-10-424-599-257786
12	6	10.3	12	9	US-09-927-436-1
13	6	10.3	12	14	US-10-161-791-268
14	6	10.3	47	12	US-10-210-172-170
15	6	10.3	52	12	US-10-424-599-164611

Sequence 774, App  
Sequence 229298,  
Sequence 56, Appl  
Sequence 61, Appl  
Sequence 156798,  
Sequence 221250,  
Sequence 250830,  
Sequence 178932,  
Sequence 47763, A  
Sequence 225829,  
Sequence 169059,  
Sequence 266340,  
Sequence 791, App  
Sequence 224333,  
Sequence 1200, Ap  
Sequence 6237, Ap  
Sequence 234852,  
Sequence 279348,  
Sequence 197414,  
Sequence 216096,  
Sequence 158794,  
Sequence 73863, A  
Sequence 205, App  
Sequence 159623,  
Sequence 159963,  
Sequence 203486,  
Sequence 279792,  
Sequence 2076, Ap  
Sequence 200341,

6 10.3 56 11 US-09-864-408A-774  
6 10.3 62 12 US-10-424-599-229298  
6 10.3 63 14 US-10-125-258-56  
6 10.3 63 14 US-10-125-258-61  
6 10.3 71 12 US-10-424-599-156798  
6 10.3 71 12 US-10-424-599-221250  
6 10.3 71 12 US-10-424-599-250830  
6 10.3 72 16 US-10-437-963-178932  
6 10.3 75 9 US-09-864-761-47763  
6 10.3 81 12 US-10-424-599-225829  
6 10.3 96 12 US-10-424-599-169059  
6 10.3 99 12 US-10-424-599-266340  
6 10.3 106 11 US-09-833-445-791  
6 10.3 107 12 US-10-424-599-224333  
6 10.3 107 15 US-10-289-762-1200  
6 10.3 108 9 US-09-738-626-6237  
6 10.3 110 12 US-10-424-599-234852  
6 10.3 113 12 US-10-424-599-279348  
6 10.3 114 12 US-10-424-599-197414  
6 10.3 117 12 US-10-424-599-216096  
6 10.3 119 12 US-10-424-599-145916  
6 10.3 125 16 US-10-437-963-158794  
6 10.3 126 12 US-10-282-122A-73863  
6 10.3 130 13 US-10-001-857-205  
6 10.3 144 12 US-10-424-599-159623  
6 10.3 153 12 US-10-424-599-159963  
6 10.3 161 12 US-10-424-599-203486  
6 10.3 161 12 US-10-424-599-279792  
6 10.3 161 16 US-10-408-765A-2076  
6 10.3 161 16 US-10-437-963-200341

#### ALIGNMENTS

#### RESULT 1

US-10-079-754A-7  
; Sequence 7, Application US/10079754A  
; Publication No. US20020164625A1  
; GENERAL INFORMATION:  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Grigor, Murray R.  
; APPLICANT: Molenaar, Adrian J.  
; APPLICANT: Davis, Stephen R.  
; TITLE OF INVENTION: Compositions Isolated from Bovine  
; TITLE OF INVENTION: Mammary Gland and Methods for Their Use  
; FILE REFERENCE: 11000.1068  
; CURRENT APPLICATION NUMBER: US/10/079,754A  
; CURRENT FILING DATE: 2002-02-19  
; PRIOR APPLICATION NUMBER: US 09/699,146  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: US 60,162,701  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR APPLICATION NUMBER: US 09/644,190  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: US 60,150,330  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Bovine  
US-10-079-754A-7

Query Match 100.0%; Score 58; DB 13; Length 58;  
Best Local Similarity 100.0%; Pred. No. 6.9e-53;  
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MKKIFVFMALILAMIRADSEKHKRKKHGRGYFQQYQYQRYPLNYPAYPP 58  
Db 1 MKKIFVFMALILAMIRADSEKHKRKKHGRGYFQQYQYQRYPLNYPAYPP 58

```

RESULT 2
US-10-079-754A-10
; Sequence 10, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; TITLE OF INVENTION: Mammary Gland and Methods for Their Use
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60,150,330
; PRIOR FILING DATE: 2000-08-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-10

Query Match      100.0%; Score 58; DB 13; Length 58;
Best Local Similarity 100.0%; Pred. No. 6.9e-53;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKHHRGYFQQYQYQRYPLNYPAYPPP 58
DB 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKHHRGYFQQYQYQRYPLNYPAYPPP 58

RESULT 3
US-10-079-754A-8
; Sequence 8, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; TITLE OF INVENTION: Mammary Gland and Methods for Their Use
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60,150,330
; PRIOR FILING DATE: 2000-08-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-8

Query Match      87.9%; Score 51; DB 13; Length 58;
Best Local Similarity 100.0%; Pred. No. 1.4e-45;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKHHRGYFQQYQYQRYPLNYPAYPPP 58
DB 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKHHRGYFQQYQYQRYPLNYPAYPPP 58

RESULT 4
US-10-079-754A-12
; Sequence 12, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; TITLE OF INVENTION: Mammary Gland and Methods for Their Use
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-12

Query Match      58.6%; Score 34; DB 13; Length 59;
Best Local Similarity 100.0%; Pred. No. 7.2e-28;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKHH 34
DB 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKHH 34

RESULT 5
US-10-079-754A-9
; Sequence 9, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; TITLE OF INVENTION: Mammary Gland and Methods for Their Use
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-9

Query Match      87.9%; Score 51; DB 13; Length 58;
Best Local Similarity 100.0%; Pred. No. 1.4e-45;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKHHRGYFQQYQYQRYPLNYPAYPPP 58
DB 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKHHRGYFQQYQYQRYPLNYPAYPPP 58

```





```

; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David
; APPLICANT: Alsbrook II, John
; APPLICANT: Lepley, Denise et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; CURRENT APPLICATION NUMBER: US/10/210,172
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/323,994
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/373,814
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/310,544
; PRIOR FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/313,201
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/312,892
; PRIOR FILING DATE: 2001-08-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 327
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 168
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-172-168

Query Match      12.1%; Score 7; DB 12; Length 62;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19 ADSSEK 25
DB      19 ADSSEK 25
|||||

RESULT 10
US-10-424-599-145633
; Sequence 145633, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 145633
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_102525C.1.pep
US-10-424-599-145633

Query Match      12.1%; Score 7; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19 ADSSEK 25
DB      19 ADSSEK 25
|||||

RESULT 11
US-10-424-599-257786
; Sequence 257786, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 257786
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(150)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_74804C.1.pep
US-10-424-599-257786

Query Match      12.1%; Score 7; DB 12; Length 150;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      14 LAMIRAD 20
DB      20 LAMIRAD 26
|||||

RESULT 12
US-09-927-436-1
; Sequence 1, Application US/09927436
; Patent No. US20020155455A1
; GENERAL INFORMATION:
; APPLICANT: Tadavoni-Rebek, Mitra
; APPLICANT: Ameshey, Joseph W.
; APPLICANT: Rooney, Regina
; TITLE OF INVENTION: Highly Homogeneous Molecular Markers for Electrophoresis
; FILE REFERENCE: 0942.5300001
; CURRENT APPLICATION NUMBER: US/09/927,436
; CURRENT FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: US 60/224,345
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Synthetic peptide
; NAME/KEY: MOD_RES
; LOCATION: (2)..(2)
; OTHER INFORMATION: Modified with fluorescein
; NAME/KEY: MOD_RES
; LOCATION: (5)..(5)
; OTHER INFORMATION: Modified with fluorescein
US-09-927-436-1

```

Query Match 10.3%; Score 6; DB 9; Length 12;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 KRKKHH 34  
| | | | |  
DB 3 KRKKHH 8

RESULT 13

US-10-161-791-268  
; Sequence 268, Application US/10161791  
; Publication No. US20030186863A1  
; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.  
; APPLICANT: KAY, Brian K.  
; APPLICANT: THORN, Judith M.  
; APPLICANT: QUILLIAM, Lawrence A.  
; APPLICANT: DER, Channing J.  
; APPLICANT: FOWLKES, Dana M.  
; APPLICANT: RIDER, James E.  
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
; TITLE OF INVENTION: ISOLATING AND USING SAME  
; NUMBER OF SEQUENCES: 467  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/161,791  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/602,999  
; FILING DATE: 16-FEB-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-202  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 268:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-10-161-791-268

Query Match 10.3%; Score 6; DB 14; Length 12;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 YPPAYP 56  
| | | | |  
DB 1 YPPAYP 6

RESULT 14

US-10-210-172-170  
; Sequence 170, Application US/10210172  
; Publication No. US20040043928A1

GENERAL INFORMATION:

; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Miller, Charles  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Pena, Carol  
; APPLICANT: Rieger, Daniel  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Li, Li  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Casman, Stacie  
; APPLICANT: Voss, Edward  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Gorman, Linda  
; APPLICANT: Leite, Mario  
; APPLICANT: Vernet, Corine  
; APPLICANT: Anderson, David  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Zhong, Mei  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Hjalte, Tord  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Edinger, Shlomit  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel  
; APPLICANT: MacDougall, John  
; APPLICANT: Stone, David  
; APPLICANT: Alsbrook II, John  
; APPLICANT: Lepley, Denise et al.  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS  
; FILE REFERENCE: 21402-416 A  
; CURRENT APPLICATION NUMBER: US/10/210,172  
; CURRENT FILING DATE: 2001-08-01  
; PRIOR APPLICATION NUMBER: 60/309,501  
; PRIOR FILING DATE: 2001-08-02  
; PRIOR APPLICATION NUMBER: 60/323,994  
; PRIOR FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 60/373,814  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/310,291  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: 60/310,951  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: 60/310,544  
; PRIOR FILING DATE: 2001-08-07  
; PRIOR APPLICATION NUMBER: 60/311,292  
; PRIOR FILING DATE: 2001-08-09  
; PRIOR APPLICATION NUMBER: 60/311,979  
; PRIOR FILING DATE: 2001-08-13  
; PRIOR APPLICATION NUMBER: 60/313,201  
; PRIOR FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: 60/312,892  
; PRIOR FILING DATE: 2001-08-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 327  
; SOFTWARE: Curasequid version 0.1  
; SEQ ID NO 170  
; LENGTH: 47  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-210-172-170

Query Match 10.3%; Score 6; DB 12; Length 47;  
Best Local Similarity 100.0%; Pred. No. 89;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 DSSEK 25  
| | | | |  
DB 3 DSSEK 8

```
RESULT 15
US-10-424-599-164611
; Sequence 164611, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 164611
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(52)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_119661C.1.pep
US-10-424-599-164611

Query Match      10.3%; Score 6; DB 12; Length 52;
Best Local Similarity 100.0%; Pred.No. 97;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      43 PYQRYP 48
Db      35 PYQRYP 40

Search completed: August 6, 2004, 17:15:03
Job time : 47 secs
```



RESULT 2  
US-08-308-883-2  
; Sequence 2, Application US/08308883  
; Patent No. 5576300  
; GENERAL INFORMATION:  
; APPLICANT: Mukerji, P.  
; APPLICANT: Prieto, P. A.  
; APPLICANT: Seo, A. E.-Y.  
; APPLICANT: Baxter, J. H.  
; APPLICANT: Cummings, R. D.  
; TITLE OF INVENTION: Method for Inhibition of Human Rotavirus Infection.  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lonnie R. Drayer  
; ADDRESSEE: ROSS Products Division  
; ADDRESSEE: Abbott Laboratories  
; STREET: 625 Cleveland Avenue  
; CITY: Columbus  
; STATE: Ohio  
; COUNTRY: United States  
; ZIP: 43215  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb storage  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh System 7.1  
; SOFTWARE: ClarisWorks 1.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/308,883  
; FILING DATE: 16-SEP-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION INFORMATION:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (614) 624-3774  
; TELEFAX: (614) 624-3074  
; TELEX: No. 5576300e  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 182 amino acids  
; TYPE: Amino acid  
; STRANDEDNESS:  
; TOPOLOGY: Linear  
; MOLECULE TYPE: Protein.  
; DESCRIPTION:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION:  
; PUBLICATION INFORMATION:  
; AUTHORS: L. Hansson et al  
; TITLE: DNA Encoding Kappa-Casein, Process for Obtaining the Protein and Use The  
; JOURNAL:

; VOLUME:  
; ISSUE:  
; PAGES:  
; DATE:  
; DOCUMENT NUMBER: PCT/WO93/15196  
; FILING DATE: 25-JAN-1993  
; PUBLICATION DATE: 05-AUG-1993  
; RELEVANT RESIDUES IN SEQ ID NO:  
US-08-308-883-2  
Query Match 19.9%; Score 63; DB 1; Length 182;  
Best Local Similarity 28.8%; Pred. No. 2.3;  
Matches 17; Conservative 13; Mismatches 25; Indels 4; Gaps 3;  
QY 1 MKRIFVF-INALILAMRADSSSEKRRKKGHRGCFQYQYQYPLNY-PPAYPF 57  
DB 1 MKSFLVVNALTLPLFLAVEVQKQKQACHENDERPFYQRTAPY--VPMYYVPNSIP 57  
RESULT 3  
US-08-730-163-2  
; Sequence 2, Application US/08730163  
; Patent No. 5712250  
; GENERAL INFORMATION:  
; APPLICANT: Mukerji, P.  
; APPLICANT: Prieto, P. A.  
; APPLICANT: Seo, A. E.-Y.  
; APPLICANT: Baxter, J. H.  
; APPLICANT: Cummings, R. D.  
; TITLE OF INVENTION: Product for Inhibition of Human Rotavirus Infection.  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lonnie R. Drayer  
; ADDRESSEE: ROSS Products Division  
; ADDRESSEE: Abbott Laboratories  
; STREET: 625 Cleveland Avenue  
; CITY: Columbus  
; STATE: Ohio  
; COUNTRY: United States  
; ZIP: 43215  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb storage (B) COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh System 7.1 (D) SOFTWARE: ClarisWorks 1.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/730,163  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/308,882  
; FILING DATE: 16-SEP-1994  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (614) 624-3774  
; TELEFAX: (614) 624-3074  
; TELEX: No. 5712250e  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 182 amino acids  
; TYPE: Amino acid  
; STRANDEDNESS:  
; TOPOLOGY: Linear  
; MOLECULE TYPE: Protein.  
; DESCRIPTION:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:







```

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21881
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21881

Query Match      18.7%; Score 59; DB 4; Length 720;
Best Local Similarity 29.8%; Pred. No. 32;
Matches 14; Conservative 8; Mismatches 15; Indels 10; Gaps 1;

QY      18 RADSSSEKHKRKKKH-----RGYFOQYQYQYPLNPPA 54
      |||:||||:|||||
Db      325 RAGQOPQRHRQRRHPPGGRPAGORPAGTGTGQGPARRQPRRPA 371

RESULT 13
US-09-252-991A-27419
; Sequence 27419, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27419
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27419

Query Match      18.5%; Score 58.5; DB 4; Length 247;
Best Local Similarity 39.4%; Pred. No. 11;
Matches 13; Conservative 6; Mismatches 11; Indels 3; Gaps 1;

QY      18 RADSSSEKHKRKKKHGCGYFOQYQYQYPLN 50
      |||:||||:|||||
Db      60 RRDTGGRHPRRLDHRHRRQQQPP---HPLN 89

RESULT 14
US-09-252-991A-20994
; Sequence 20994, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20994
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20994

```

```

Query Match      18.5%; Score 58.5; DB 4; Length 333;
Best Local Similarity 38.7%; Pred. No. 16;
Matches 12; Conservative 5; Mismatches 11; Indels 3; Gaps 1;

QY      25 KHRHKK---KHHRGYFQQYQPYQRYPLNYP 52
DB      114 RRHHRRPDQGTHRRGLRQGLPVRRLFADHP 144

RESULT 15
US-09-252-991A-30311
; Sequence 30311, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30311
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30311

```

```

Query Match      18.4%; Score 58; DB 4; Length 521;
Best Local Similarity 36.7%; Pred. NC. 30;
Matches 18; Conservative 6; Mismatches 21; Indels 4; Gaps 2;

QY      12 LILAMRADGSE--EKSRKRKKHRRGYQQQYQYQYVPLNYPAYPPF 58
DB      45 LIRTTIAQGVIEVHRHRRRLHUHQHPLQRRLQRLQ--RRFAHQRP 91

```

Search completed: August 6, 2004, 17:05:15  
Job time : 19 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2004, 17:04:44 ; Search time 47 Seconds  
(without alignments)

387.098 Million cell updates/sec

Title: US-10-079-754A-10

Perfect score: 316

Sequence: 1 MKKIFVFMALILAMIRAD.....QQYQYQRYPLNYPPAYPPF 58

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1291235 seqs, 313682936 residues

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpa/PCF\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpa/PCFUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	316	100.0	58	US-10-079-754A-7	Sequence 7, Appli
2	316	100.0	58	US-10-079-754A-10	Sequence 10, Appli
3	315	99.7	58	US-10-079-754A-8	Sequence 8, Appli
4	170	53.8	59	US-10-079-754A-12	Sequence 12, Appli
5	169	53.5	70	US-10-079-754A-9	Sequence 9, Appli
6	112	35.4	21	US-10-079-754A-11	Sequence 11, Appli
7	110	34.8	62	US-10-210-172-168	Sequence 168, App
8	95	30.1	51	US-09-917-340-19	Sequence 19, Appli
9	94.5	29.9	78	US-09-992-600A-6	Sequence 6, Appli
10	94.5	29.9	78	US-09-924-340-6	Sequence 6, Appli
11	94.5	29.9	78	US-09-992-095B-6	Sequence 6, Appli
12	94.5	29.9	78	US-09-999-570-6	Sequence 6, Appli
13	94.5	29.9	78	US-10-000-489-6	Sequence 6, Appli
14	94.5	29.9	78	US-10-000-986-6	Sequence 6, Appli
15	94.5	29.9	78	US-10-154-678-6	Sequence 6, Appli

16	22.8	1504	16	US-10-437-963-192104	Sequence 192104,
17	22.8	1680	16	US-10-437-963-192175	Sequence 192175,
18	22.2	1363	16	US-10-437-963-192041	Sequence 192041,
19	21.8	229	12	US-10-282-122A-71761	Sequence 71761, A
20	21.5	1822	16	US-10-437-963-192099	Sequence 192099,
21	20.9	277	16	US-10-631-581-22	Sequence 22, Appli
22	20.6	1081	16	US-10-437-963-191989	Sequence 191989,
23	20.6	1199	16	US-10-437-963-192108	Sequence 192108,
24	20.6	1403	16	US-10-437-963-191987	Sequence 191987,
25	20.6	1487	16	US-10-437-963-188109	Sequence 188109,
26	20.6	1504	15	US-10-374-780A-1457	Sequence 1457, Ap
27	20.6	1504	16	US-10-437-963-192134	Sequence 192134,
28	20.6	1516	16	US-10-437-963-192071	Sequence 192071,
29	20.6	1542	16	US-10-437-963-192098	Sequence 192098,
30	20.6	1562	15	US-10-374-780A-1458	Sequence 1458, Ap
31	20.6	1562	15	US-10-374-780A-1459	Sequence 1459, Ap
32	20.6	1571	16	US-10-437-963-192152	Sequence 192152,
33	20.6	1572	16	US-10-437-963-192213	Sequence 192213,
34	20.6	1595	16	US-10-437-963-179538	Sequence 179538,
35	20.6	1629	16	US-10-437-963-192058	Sequence 192058,
36	20.6	1681	16	US-10-437-963-191994	Sequence 191994,
37	20.6	1694	16	US-10-437-963-192057	Sequence 192057,
38	20.6	1732	16	US-10-437-963-192167	Sequence 192167,
39	20.6	1761	16	US-10-437-963-192096	Sequence 192096,
40	20.6	1764	16	US-10-437-963-192101	Sequence 192101,
41	20.6	1776	16	US-10-437-963-179541	Sequence 179541,
42	20.6	1792	16	US-10-437-963-192174	Sequence 192174,
43	20.6	1809	16	US-10-437-963-192109	Sequence 192109,
44	20.6	1815	16	US-10-437-963-192168	Sequence 192168,
45	20.6	1820	16	US-10-437-963-192165	Sequence 192165,

ALIGNMENTS

RESULT 1

US-10-079-754A-7  
; Sequence 7, Application US/10079754A  
; Publication No. US20020164625A1  
; GENERAL INFORMATION:  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Grigor, Murray R.  
; APPLICANT: Molenaar, Adrian J.  
; APPLICANT: Davis, Stephen R.  
; TITLE OF INVENTION: Compositions Isolated from Bovine  
; TITLE OF INVENTION: Mammary Gland and Methods for Their Use  
; FILE REFERENCE: 11000.1068  
; CURRENT APPLICATION NUMBER: US/10/079,754A  
; CURRENT FILING DATE: 2002-02-19  
; PRIOR APPLICATION NUMBER: US 09/699,146  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: US 60,162,701  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR APPLICATION NUMBER: US 09/644,190  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: US 60,150,330  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Bovine  
US-10-079-754A-7

Query Match 100.0%; Score 316; DB 13; Length 58;

Best Local Similarity 100.0%; Pred. No. 5.6e-30; Indels 0; Gaps 0;  
Matches 58; Conservative 0; Mismatches 0;

QY 1 MKKIFVFMALILAMIRADSEKRRKXKHHRGYFQQYQYQRYPLNYPPAYPPF 58

Db 1 MKKIFVFMALILAMIRADSEKRRKXKHHRGYFQQYQYQRYPLNYPPAYPPF 58

```

RESULT 2
US-10-079-754A-10
; Sequence 10, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-10

Query Match      100.0%; Score 316; DB 13; Length 58;
Best Local Similarity 100.0%; Pred. No. 5.6e-30;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKHHRGYFQQYQPYQRYPLNYPAYPPP 58
DB 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKHHRGYFQQYQPYQRYPLNYPAYPPP 58

RESULT 3
US-10-079-754A-8
; Sequence 8, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-8

Query Match      99.7%; Score 315; DB 13; Length 58;
Best Local Similarity 98.3%; Pred. No. 7.4e-30;
Matches 57; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKHHRGYFQQYQPYQRYPLNYPAYPPP 58
DB 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKHHRGYFQQYQPYQRYPLNYPAYPPP 58

RESULT 4
US-10-079-754A-12
; Sequence 12, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-12

Query Match      53.8%; Score 170; DB 13; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKHH 34
DB 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKHH 34

RESULT 5
US-10-079-754A-9
; Sequence 9, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-9

Query Match      99.7%; Score 315; DB 13; Length 58;
Best Local Similarity 98.3%; Pred. No. 7.4e-30;
Matches 57; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```
US-10-079-754A-9
Query Match      53.5%; Score 169; DB 13; Length 70;
Best Local Similarity 97.1%; Pred. No. 1.7e-12;
Matches 33; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKIFIFVIMAILAMIRADSSSEKHKRKKKH 34
Db 1 MKIFIFVIMAILAMIRADSSSEKHKRKKKH 34

RESULT 6
US-10-079-754A-11
; Sequence 11, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; TITLE OF INVENTION: Mammary Gland and Methods for Their Use
; FILE REFERENCE: 11000.11088
; CURRENT APPLICATION NUMBER: US/10/079,754A
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-11

Query Match      35.4%; Score 112; DB 13; Length 21;
Best Local Similarity 94.7%; Pred. No. 2.5e-06;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 38 FQQQPYQRYPLNYPYPAYP 56
Db 1 FQQQPYQRYPLNYPYPAYP 19

RESULT 7
US-10-210-172-168
; Sequence 168, Application US/10210172
; Publication No. US20040043928A1
; GENERAL INFORMATION:
; APPLICANT: Rekuda, Ramesh
; APPLICANT: Miller, Charles
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol
; APPLICANT: Rieger, Daniel
; APPLICANT: Shimkets, Richard
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Li, Li
; APPLICANT: Ji, Weizhen
; APPLICANT: Padigaru, Murallidhara
; APPLICANT: Casman, Stacie
; APPLICANT: Voss, Edward
; APPLICANT: Boldog, Ferenc
; APPLICANT: Gorman, Linda
; APPLICANT: Leite, Mario
; APPLICANT: Vernet, Corine
; APPLICANT: Anderson, David
; APPLICANT: Guo, Xiaojia

US-10-079-754A-10.rapb
; APPLICANT: Zhong, Mei
; APPLICANT: Gerlach, Valerie
; APPLICANT: Hjalte, Iord
; APPLICANT: Rastelli, Luca
; APPLICANT: Spytek, Kimberly
; APPLICANT: Edinger, Shalomit
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David
; APPLICANT: Alsbrook II, John
; APPLICANT: Lepley, Denise et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-416 A
; CURRENT APPLICATION NUMBER: US/10/210,172
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/323,994
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/373,814
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/310,544
; PRIOR FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/313,201
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/312,892
; PRIOR FILING DATE: 2001-08-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 327
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 168
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-172-168

Query Match      34.8%; Score 110; DB 12; Length 62;
Best Local Similarity 49.1%; Pred. No. 1.4e-05;
Matches 27; Conservative 7; Mismatches 19; Indels 2; Gaps 2;

Qy 1 MKIFIFVIMAILAMIRADSSSEKHKRKKKHRRGYFQQYQYQRYPLNYPAYP 55
Db 1 MKFLVFAFILALVSMIGADSSSEKFLRRIGRFGYGY-CPYQVPEQPL-YQPY 53

RESULT 8
US-09-917-340-19
; Sequence 19, Application US/09917340
; Patent No. US20020090369A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Christopher J.
; APPLICANT: McNulty, Jonathan F.
; APPLICANT: Reid, Ted W.
; TITLE OF INVENTION: Transplant Media
; FILE REFERENCE: TPLANT-06468
; CURRENT APPLICATION NUMBER: US/09/917,340
; CURRENT FILING DATE: 2001-07-29
; PRIOR APPLICATION NUMBER: 60/221,632
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/249,602
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/290,932
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 96
```

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 19

; LENGTH: 51

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-917-340-19

Query Match 30.1%; Score 95; DB 9; Length 51;  
Best Local Similarity 46.8%; Pred. No. 0.00067;  
Matches 22; Conservative 6; Mismatches 13; Indels 7; Gaps 2;

QY 1 MKIFIFVFMALILAMIRADSSSEKRRH-RKKKH-----HRGYFQQY 41  
DB 1 MKFFVFALVLMISMTGADSHAKRHGGRKFKHEKHHSHRGYSNY 47

RESULT 9

US-09-992-600A-6

; Sequence 6, Application US/09992600A

; Publication No. US20030027161A1

; GENERAL INFORMATION:

; APPLICANT: Benjamin, Stephanie

; APPLICANT: Tanaka, Hiroaki

; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF

; FILE REFERENCE: 91 US4, DIV

; CURRENT APPLICATION NUMBER: US/09/992,600A

; CURRENT FILING DATE: 2001-11-13

; PRIOR APPLICATION NUMBER: US/09/924,340

; PRIOR FILING DATE: 2001-08-06

; PRIOR APPLICATION NUMBER: PCT/IB01/01715

; PRIOR FILING DATE: 2001-08-06

; PRIOR APPLICATION NUMBER: US 60/305,456

; PRIOR FILING DATE: 2001-07-13

; PRIOR APPLICATION NUMBER: US 60/302,277

; PRIOR FILING DATE: 2001-06-29

; PRIOR APPLICATION NUMBER: US 60/298,698

; PRIOR FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: US 60/293,574

; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 114

; SOFTWARE: JPatent

; SEQ ID NO 6

; LENGTH: 78

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SIGNAL

; LOCATION: 1..19

US-09-992-600A-6

Query Match 29.9%; Score 94.5; DB 10; Length 78;  
Best Local Similarity 32.7%; Pred. No. 0.0012;  
Matches 18; Conservative 13; Mismatches 17; Indels 7; Gaps 1;

QY 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKKHGRGYFQQYQRYPLNYPAY 55  
DB 1 MKFFVFALVLMISMTGADSHAKRHGGRKFKH-----EKHSHYHITLLPLF 48

RESULT 10

US-09-924-340-6

; Sequence 6, Application US/09924340

; Publication No. US20030027248A1

; GENERAL INFORMATION:

; APPLICANT: Benjamin, Stephanie

; APPLICANT: Tanaka, Hiroaki

; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF

; FILE REFERENCE: 91 US2, REG

; CURRENT APPLICATION NUMBER: US/09/924,340

; CURRENT FILING DATE: 2001-08-06

; PRIOR APPLICATION NUMBER: US 60/305,456

; PRIOR FILING DATE: 2001-07-13

; PRIOR APPLICATION NUMBER: US 60/302,277

; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPatent  
; SEQ ID NO 6  
; LENGTH: 78  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: 1..19  
; US-09-924-340-6

Query Match 29.9%; Score 94.5; DB 10; Length 78;  
Best Local Similarity 32.7%; Pred. No. 0.0012;  
Matches 18; Conservative 13; Mismatches 17; Indels 7; Gaps 1;

QY 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKKHGRGYFQQYQRYPLNYPAY 55  
DB 1 MKFFVFALVLMISMTGADSHAKRHGGRKFKH-----EKHSHYHITLLPLF 48

RESULT 11

US-09-992-095B-6

; Sequence 6, Application US/09992095B

; Publication No. US20030157485A1

; GENERAL INFORMATION:

; APPLICANT: Benjamin, Stephanie

; APPLICANT: Tanaka, Hiroaki

; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF

; FILE REFERENCE: 91 US5, DIV

; CURRENT APPLICATION NUMBER: US/09/992,095B

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: US 09/924,340

; PRIOR FILING DATE: 2001-08-06

; PRIOR APPLICATION NUMBER: PCT/IB01/01715

; PRIOR FILING DATE: 2001-08-06

; PRIOR APPLICATION NUMBER: US 60/305,456

; PRIOR FILING DATE: 2001-07-13

; PRIOR APPLICATION NUMBER: US 60/302,277

; PRIOR FILING DATE: 2001-06-29

; PRIOR APPLICATION NUMBER: US 60/298,698

; PRIOR FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: US 60/293,574

; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 112

; SOFTWARE: JPatent

; SEQ ID NO 6

; LENGTH: 78

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SIGNAL

; LOCATION: 1..19

US-09-992-095B-6

Query Match 29.9%; Score 94.5; DB 10; Length 78;  
Best Local Similarity 32.7%; Pred. No. 0.0012;  
Matches 18; Conservative 13; Mismatches 17; Indels 7; Gaps 1;

QY 1 MKIFIFVFMALILAMIRADSSSEKRRKRRKKHGRGYFQQYQRYPLNYPAY 55  
DB 1 MKFFVFALVLMISMTGADSHAKRHGGRKFKH-----EKHSHYHITLLPLF 48

RESULT 12

US-09-999-570-6

; Sequence 6, Application US/09999570

; Publication No. US20030170628A1

; GENERAL INFORMATION:

APPLICANT: Benjanin, Stephane  
APPLICANT: Tanaka, Hiroaki  
TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
FILE REFERENCE: G-091US08DIV  
CURRENT APPLICATION NUMBER: US/09/999,570  
CURRENT FILING DATE: 2001-06-14  
PRIOR APPLICATION NUMBER: US 09/924,340  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: PCT/IB01/01715  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: US 60/305,456  
PRIOR FILING DATE: 2001-07-13  
PRIOR APPLICATION NUMBER: US 60/302,277  
PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: US 60/298,698  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 60/293,574  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 112  
SOFTWARE: JPatent  
SEQ ID NO 6  
LENGTH: 78  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: 1..19  
US-09-999-570-6

Query Match 29.9%; Score 94.5; DB 10; Length 78;  
Best Local Similarity 32.7%; Pred. No. 0.0012;  
Matches 18; Conservative 13; Mismatches 17; Indels 7; Gaps 1;

QY 1 MKFIFVFMALILAMIRADSESEKRRKRRKKGHRGYFQQYQYQYPLNYPAY 55  
DB 1 MKFFVFLVLMISMISADSHKRRHGYRRKFH-----EKHSYHITLLPLF 48

RESULT 13

US-10-000-489-6  
Sequence 6, Application US/10000489  
Publication No. US2003009201A1  
GENERAL INFORMATION:  
APPLICANT: Benjanin, Stephane  
APPLICANT: Tanaka, Hiroaki  
TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
FILE REFERENCE: 91.US6.DIV  
CURRENT APPLICATION NUMBER: US/10/000,489  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: US 09/924,340  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: PCT/IB01/01715  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: US 60/305,456  
PRIOR FILING DATE: 2001-07-13  
PRIOR APPLICATION NUMBER: US 60/302,277  
PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: US 60/298,698  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 60/293,574  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 112  
SOFTWARE: JPatent  
SEQ ID NO 6  
LENGTH: 78  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: 1..19  
US-10-000-489-6

Query Match 29.9%; Score 94.5; DB 14; Length 78;

Best Local Similarity 32.7%; Pred. No. 0.0012;  
Matches 18; Conservative 13; Mismatches 17; Indels 7; Gaps 1;  
QY 1 MKFIFVFMALILAMIRADSESEKRRKRRKKGHRGYFQQYQYQYPLNYPAY 55  
DB 1 MKFFVFLVLMISMISADSHKRRHGYRRKFH-----EKHSYHITLLPLF 48

RESULT 14

US-10-000-986-6  
Sequence 6, Application US/10000986  
Publication No. US20030096247A1  
GENERAL INFORMATION:  
APPLICANT: Benjanin, Stephane  
APPLICANT: Tanaka, Hiroaki  
TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
FILE REFERENCE: 91.US9.DIV  
CURRENT APPLICATION NUMBER: US/10/000,986  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: US 09/924,340  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: PCT/IB01/01715  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: US 60/305,456  
PRIOR FILING DATE: 2001-07-13  
PRIOR APPLICATION NUMBER: US 60/302,277  
PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: US 60/298,698  
PRIOR FILING DATE: 2001-08-15  
PRIOR APPLICATION NUMBER: US 60/293,574  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 112  
SOFTWARE: JPatent  
SEQ ID NO 6  
LENGTH: 78  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: 1..19  
US-10-000-986-6

Query Match 29.9%; Score 94.5; DB 14; Length 78;  
Best Local Similarity 32.7%; Pred. No. 0.0012;  
Matches 18; Conservative 13; Mismatches 17; Indels 7; Gaps 1;

QY 1 MKFIFVFMALILAMIRADSESEKRRKRRKKGHRGYFQQYQYQYPLNYPAY 55  
DB 1 MKFFVFLVLMISMISADSHKRRHGYRRKFH-----EKHSYHITLLPLF 48

RESULT 15

US-10-154-678-6  
Sequence 6, Application US/10154678  
Publication No. US20030162186A1  
GENERAL INFORMATION:  
APPLICANT: Benjanin, Stephane  
APPLICANT: Tanaka, Hiroaki  
TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
FILE REFERENCE: 182.US1.REG  
CURRENT APPLICATION NUMBER: US/10/154,678  
CURRENT FILING DATE: 2002-10-15  
PRIOR APPLICATION NUMBER: US 09/924,340  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: US 60/305,456  
PRIOR FILING DATE: 2001-07-13  
PRIOR APPLICATION NUMBER: US 60/302,277  
PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: US 60/298,698  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 60/293,574  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 112





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 16, 2004, 00:50:06 ; Search time 61 Seconds  
(without alignments)  
527.658 Million cell updates/sec

Title: US-10-079-754A-10

Perfect score: 58

Sequence: 1 MKIFIFVIMALILAMTRAD.....QQYQYQRYPLNPPAYPPF 58

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1360396

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US10079754/runat\_06082004\_181355\_29928/app\_query.fasta\_1.199  
-DB=Issued.Patents NA -QMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTPMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10079754@cgn\_1\_1\_65@runat\_06082004\_181355\_29928 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORE=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	13.8	1438	3	US-08-434-099A-26
2	8	13.8	10564	1	US-08-206-176-5
3	7	12.1	711	4	US-09-543-681A-2950
4	7	12.1	1062	4	US-09-543-681A-1164
C 5	7	12.1	1509	4	US-09-134-000C-1868
C 6	7	12.1	1914	4	US-09-543-681A-2340
C 7	7	12.1	2082	4	US-09-543-681A-2670
C 8	7	12.1	2437	1	US-08-456-647B-3
C 9	7	12.1	2437	2	US-08-237-401A-3
C 10	7	12.1	2598	1	US-08-100-692-4
C 11	7	12.1	2598	2	US-08-674-030-4
C 12	7	12.1	2619	4	US-09-166-350-28
					Sequence 26, Appl
					Sequence 5, Appl
					Sequence 2950, Ap
					Sequence 1164, Ap
					Sequence 1868, Ap
					Sequence 2340, Ap
					Sequence 2670, Ap
					Sequence 3, Appl
					Sequence 3, Appl
					Sequence 4, Appl
					Sequence 4, Appl
					Sequence 28, Appl

Sequence 17, Appl	7	12.1	2624	2	US-08-486-663A-17
Sequence 7, Appl	7	12.1	2624	3	US-08-247-904B-7
Sequence 20, Appl	7	12.1	2624	3	US-08-767-942A-20
Sequence 1, Appl	7	12.1	2625	6	US-08-435-637A-1
Patent No. 5198347	7	12.1	3157	6	5198347-3
Sequence 5, Appl	7	12.1	3321	1	US-08-484-438-5
Sequence 3, Appl	7	12.1	5484	3	US-09-633-580A-3
Sequence 1, Appl	7	12.1	5501	1	US-08-484-438-1
Sequence 3, Appl	7	12.1	5555	1	US-08-484-438-3
Sequence 37, Appl	7	12.1	9834	4	US-08-956-171E-37
Sequence 1, Appl	7	12.1	1830121	4	US-09-557-884-1
Sequence 1, Appl	7	12.1	1830121	4	US-09-643-990A-1
Sequence 30, Appl	6	10.3	38	1	US-07-967-693-30
Sequence 28, Appl	6	10.3	38	1	US-08-195-072-28
Sequence 28, Appl	6	10.3	38	1	US-08-195-735-28
Sequence 28, Appl	6	10.3	38	1	US-08-195-747-28
Sequence 28, Appl	6	10.3	38	1	US-08-446-884-28
Sequence 28, Appl	6	10.3	38	1	US-08-195-073-28
Sequence 28, Appl	6	10.3	38	1	US-08-198-175-28
Sequence 28, Appl	6	10.3	38	2	US-08-443-153-28
Sequence 28, Appl	6	10.3	38	3	US-08-442-807-28
Sequence 17, Appl	6	10.3	47	3	US-08-856-331-17
Sequence 71, Appl	6	10.3	51	2	US-08-477-527A-71
Sequence 71, Appl	6	10.3	51	3	US-08-481-710-71
Sequence 33, Appl	6	10.3	51	5	PCT-US96-09537-71
Sequence 34, Appl	6	10.3	69	2	US-08-477-527A-33
Sequence 38, Appl	6	10.3	69	2	US-08-477-527A-36
Sequence 33, Appl	6	10.3	69	3	US-08-481-710-33
Sequence 34, Appl	6	10.3	69	3	US-08-481-710-34
Sequence 38, Appl	6	10.3	69	3	US-08-481-710-38
Sequence 33, Appl	6	10.3	69	5	PCT-US96-09537-33
Sequence 34, Appl	6	10.3	69	5	PCT-US96-09537-34

ALIGNMENTS

RESULT 1

US-08-434-099A-26  
; Sequence 26, Application US/08434099A  
; Patent No. 6083902  
; GENERAL INFORMATION:  
; APPLICANT: Cedarholm-Wms., Stewart A.  
; TITLE OF INVENTION: Recombinant Fibrin Chains,  
; TITLE OF INVENTION: Fibrin and Fibrin-Homologs  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: E.R. Squibb & Sons, Inc.  
; STREET: 100 Headquarters Park Drive  
; CITY: Skillman  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08558  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/434,099A  
; FILING DATE: 03-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/236,979  
; FILING DATE: 02-MAY-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Furman, Jr., Esq., Theodore R  
; REGISTRATION NUMBER: 30,942  
; REFERENCE/DOCKET NUMBER: CV0054a  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 908-281-2372  
; TELEFAX: 908-281-2373  
; TELEX:

```
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1438 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 3...1364
; OTHER INFORMATION:
US-08-434-099A-26

Alignment Scores:
Pred. No.: 9.99 Length: 1438
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.79% Indels: 0
DB: 3 Gaps: 0

US-10-079-754A-10 (1-58) x US-08-434-099A-26 (1-1438)
QY 10 MetAlaLeuIleLeuAlaMetIle 17
Db 1003 ATGGCTTTGATTGGCGATGATC 1026

RESULT 2
US-08-206-176-5
; Sequence 5, Application US/08206176
; Patent No. 5639940
; GENERAL INFORMATION:
; APPLICANT: Garner, Ian
; APPLICANT: Dalrymple, Michael A
; APPLICANT: Prunkard, Donna E
; APPLICANT: Foster, Donald C
; TITLE OF INVENTION: Production of Fibrinogen in Transgenic
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/206,176
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 93-15
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1056 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: human fibrinogen gamma chain
; FEATURE:
; NAME/KEY: CDS
; LOCATION: Join(1799...1876, 1973...2017, 2207...2390, 2510

; LOCATION: ..2603, 4211...4341, 4645...4778, 5758...5942, 7426
; LOCATION: ..7703, 9342...9571)
US-08-206-176-5

Alignment Scores:
Pred. No.: 63 Length: 1056
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.79% Indels: 0
DB: 1 Gaps: 0

US-10-079-754A-10 (1-58) x US-08-206-176-5 (1-1056)
QY 10 MetAlaLeuIleLeuAlaMetIle 17
Db 7524 ATGGCTTTGATTGGCGATGATC 7547

RESULT 3
US-09-543-681A-2950
; Sequence 2950, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2950
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-2950

Alignment Scores:
Pred. No.: 61.8 Length: 711
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 4 Gaps: 0

US-10-079-754A-10 (1-58) x US-09-543-681A-2950 (1-711)
QY 8 PheIleMetAlaLeuIleLeu 14
Db 377 TTCATAATGCGATTGATTTG 397

RESULT 4
US-09-543-681A-1164
; Sequence 1164, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1164
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-1164

Alignment Scores:
```

Pred. No.: 89.5 Length: 1062  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.07% Indels: 0  
DB: 4 Gaps: 0

US-10-079-754A-10 (1-58) x US-09-543-681A-1164 (1-1062)

Qy 8 PheleMetalaLeulleLeu 14  
Db 467 TTTATTATGGCACTAATCTTA 487

## RESULT 5

US-09-134-000C-1868/c  
; Sequence 1868, Application US/09134000C  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1868  
; LENGTH: 1509  
; TYPE: DNA  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-1868

Alignment Scores:  
Pred. No.: 124 Length: 1509  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.07% Indels: 0  
DB: 4 Gaps: 0

US-10-079-754A-10 (1-58) x US-09-134-000C-1868 (1-1509)

Qy 3 llePhelePheValPheille 9  
Db 1212 ATTTTATTTTGTTCATC 1192

## RESULT 6

US-09-543-681A-2340/c  
; Sequence 2340, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 2340  
; LENGTH: 1914  
; TYPE: DNA  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-2340

Alignment Scores:  
Pred. No.: 154 Length: 1914  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 12.07% Indels: 0  
DB: 4 Gaps: 0

US-10-079-754A-10 (1-58) x US-09-543-681A-2340 (1-1914)

Qy 3 llePhelePheValPheille 9  
Db 669 ATCTTCATCTCGTCTTCATC 649

## RESULT 7

US-09-543-681A-2670/c  
; Sequence 2670, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 2670  
; LENGTH: 2082  
; TYPE: DNA  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-2670

Alignment Scores:  
Pred. No.: 167 Length: 2082  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.07% Indels: 0  
DB: 4 Gaps: 0

US-10-079-754A-10 (1-58) x US-09-543-681A-2670 (1-2082)

Qy 51 TyrProProAlaTyrProPhe 57  
Db 1399 TACCCGCCAGCGTATCCATTT 1379

## RESULT 8

US-08-456-647B-3/c  
; Sequence 3, Application US/08456647B  
; Patent No. 5811516  
; GENERAL INFORMATION:  
; APPLICANT: Lemke Ph.D. et al., Greg E.  
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,647B  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/237,401  
; FILING DATE: 02-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/884,486  
; FILING DATE: 15-MAY-1992

```
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell Ph.D., John R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: 07251/007002
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2437 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
IMMEDIATE SOURCE:
CLONE: Tyro-2
FEATURE:
NAME/KEY: CDS
LOCATION: 3..2118
US-08-456-647B-3
Alignment Scores:
Pred. No.: 193
Score: 7.00
Length: 2437
Matches: 7
Percent Similarity: 100.00%
Conservative: 0
Best Local Similarity: 100.00%
Mismatch: 0
Query Match: 12.07%
Indels: 0
DB: 1
Gaps: 0
US-10-079-754A-10 (1-58) x US-08-456-647B-3 (1-2437)
QY 13 lletleulMetlleAtgAala 19
Db 565 ATACTGCCATGATCAGAGCC 545
RESULT 9
US-08-237-401A-3/c
Sequence 3, Application US/08237401A
Patent No. 5837448
GENERAL INFORMATION:
APPLICANT: Lemke Ph.D. et al., Greg E.
TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/237,401A
FILING DATE: 02-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/884,486
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Haile Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07251/007001
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2437 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
IMMEDIATE SOURCE:
CLONE: Tyro-2
FEATURE:
NAME/KEY: CDS
LOCATION: 3..2118
US-08-456-647B-3
Alignment Scores:
Pred. No.: 193
Score: 7.00
Length: 2437
Matches: 7
Percent Similarity: 100.00%
Conservative: 0
Best Local Similarity: 100.00%
Mismatch: 0
Query Match: 12.07%
Indels: 0
DB: 1
Gaps: 0
US-10-079-754A-10 (1-58) x US-08-456-647B-3 (1-2437)
QY 13 lletleulMetlleAtgAala 19
Db 565 ATACTGCCATGATCAGAGCC 545
RESULT 9
US-08-237-401A-3/c
Sequence 3, Application US/08237401A
Patent No. 5837448
GENERAL INFORMATION:
APPLICANT: Lemke Ph.D. et al., Greg E.
TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/237,401A
FILING DATE: 02-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/884,486
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Haile Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07251/007001
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2437 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
IMMEDIATE SOURCE:
CLONE: Tyro-2
FEATURE:
NAME/KEY: CDS
LOCATION: 3..2118
US-08-237-401A-3
Alignment Scores:
Pred. No.: 193
Score: 7.00
Length: 2437
Matches: 7
Percent Similarity: 100.00%
Conservative: 0
Best Local Similarity: 100.00%
Mismatch: 0
Query Match: 12.07%
Indels: 0
DB: 1
Gaps: 0
US-10-079-754A-10 (1-58) x US-08-237-401A-3 (1-2437)
QY 13 lletleulMetlleAtgAala 19
Db 565 ATACTGCCATGATCAGAGCC 545
RESULT 10
US-08-100-692-4/c
Sequence 4, Application US/08100692
Patent No. 5523448
GENERAL INFORMATION:
APPLICANT: Huibregtse, Jon M.
APPLICANT: Scheffner, Martin
APPLICANT: Howley, Peter M.
TITLE OF INVENTION: E6 ASSOCIATED PROTEIN AND METHODS OF USE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza, Steuart Street Tower
CITY: San Francisco
STATE: CA
COUNTRY: U.S.A.
ZIP: 94105-1492
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/100,692
FILING DATE: 19930730
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 15280-91
TELEPHONE: (206) 467-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2598 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: CDS
LOCATION: 1010..1321
US-08-100-692-4
Alignment Scores:
Pred. No.: 204
Score: 7.00
Length: 2598
Matches: 7
```

```
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 12.07%              Indels: 0
DB:                               Gaps: 0

US-10-079-754A-10 (1-58) x US-08-100-692-4 (1-2598)

QY      3 IlePheHellePheValPheIle 9
Db      543 ATCTTCATCTTTGTCTTCATC 523

RESULT 11
US-08-674-030-4/c
; Sequence 4, Application US/08674030
; Patent No. 5914389
; GENERAL INFORMATION:
; APPLICANT: Ruitbregetse, Jon M.
; APPLICANT: Scheffner, Martin
; APPLICANT: Howley, Peter M.
; TITLE OF INVENTION: E6 ASSOCIATED PROTEIN AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/674,030
; FILING DATE: 01-JUL-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/100,692
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 15280-91
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2598 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1010..1321
; US-08-674-030-4

Alignment Scores:
Pred. No.: 204      Length: 2598
Score: 7.00        Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 12.07%              Indels: 0
Gaps: 2

US-10-079-754A-10 (1-58) x US-08-674-030-4 (1-2598)

QY      3 IlePheHellePheValPheIle 9
Db      543 ATCTTCATCTTTGTCTTCATC 523
```

```
RESULT 12
US-09-166-350-28/c
; Sequence 28, Application US/09166350A
; Patent No. 6440663
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Chen, Yao
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Old, Lloyd
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alex
; TITLE OF INVENTION: Renal Cancer Associated Antigens and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: L0461/7051
; CURRENT APPLICATION NUMBER: US/09/166,350A
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: US 09/166,350
; EARLIER FILING DATE: 1998-10-05
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 2619
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-166-350-28

Alignment Scores:
Pred. No.: 206      Length: 2619
Score: 7.00        Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 12.07%              Indels: 0
Gaps: 4

US-10-079-754A-10 (1-58) x US-09-166-350-28 (1-2619)

QY      3 IlePheHellePheValPheIle 9
Db      564 ATCTTCATCTTTGTCTTCATC 544

RESULT 13
US-08-486-663A-17/c
; Sequence 17, Application US/08486663A
; Patent No. 5968761
; GENERAL INFORMATION:
; APPLICANT: Rolfe, Mark
; APPLICANT: Eckstein, Jens W.
; APPLICANT: Draetta, Giulio
; APPLICANT: Guillaume Cottarel
; TITLE OF INVENTION: Ubiquitin Conjugating Enzymes
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,663A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
```

```
;
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-029CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2625
US-08-486-663A-17

Alignment Scores:
Pred. No.: 206 Length: 2624
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 2 Gaps: 0

US-10-079-754A-10 (1-58) x US-08-486-663A-17 (1-2624)

QY 3 llePheHlePheValPheHle 9
Db 543 ATCTTCATCTTGTGCTTCATC 523

RESULT 14
US-08-247-904B-7/c
; Sequence 7, Application US/08247904B
; Patent No. 5981699
; GENERAL INFORMATION:
; APPLICANT: Rolfe, Mark
; APPLICANT: Eckstein, Jens W.
; APPLICANT: Draetta, Giulio
; TITLE OF INVENTION: Human Ubiquitin Conjugating Enzyme
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/247,904B
; FILING DATE: 23-MAY-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-029.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
```

```
;
; LOCATION: 1..2624
US-08-247-904B-7

Alignment Scores:
Pred. No.: 206 Length: 2624
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 2 Gaps: 0

US-10-079-754A-10 (1-58) x US-08-247-904B-7 (1-2624)

QY 3 llePheHlePheValPheHle 9
Db 543 ATCTTCATCTTGTGCTTCATC 523

RESULT 15
US-08-767-942A-20/c
; Sequence 20, Application US/08767942A
; Patent No. 8068932
; GENERAL INFORMATION:
; APPLICANT: Rolfe, Mark
; APPLICANT: Chiu, M. Isabel
; APPLICANT: Berlin, Vivian
; APPLICANT: Damagnez, Veronique
; APPLICANT: Draetta, Giulio
; APPLICANT: Guillaume, Cottarel
; TITLE OF INVENTION: UBIQUITIN CONJUGATING ENZYMES
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/767,942A
; FILING DATE: 17-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-029.04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2622
US-08-767-942A-20

Alignment Scores:
Pred. No.: 206 Length: 2624
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 3 Gaps: 0

US-10-079-754A-10 (1-58) x US-08-767-942A-20 (1-2624)
```



This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 16, 2004, 01:12:59 ; Search time 326 Seconds

(without alignments)  
872.955 Million cell updates/sec

Title: US-10-079-754A-10

Perfect score: 58

Sequence: 1 MKFIFVFMALILAMIRAD.....QQYQYQRYPLNYPYPAYPPF 58

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6442686

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US1007954/runat\_06082004\_181357\_29954/app\_query.fasta\_1.199  
-DB=Published Applications NA -OFMT=fastap -SUFFIX=rbp -MINMATCH=0.1  
-LOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR\_SCORE=quality -THR\_MIN=1  
-ALIGN=15 -MODS=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US1007954 @CGN 1 1 480 @runat\_06082004\_181357\_29954  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRADS=1 -XGAPOP=60 -XGAPEXT=60  
-Fgapop=6 -Fgapext=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
------------	-------	-------	-------	--------	-------	-------------

1	58	100.0	267	14	US-10-079-754A-15	Sequence 15, Appl
2	58	100.0	267	14	US-10-079-623-200	Sequence 200, App
3	58	100.0	505	14	US-10-079-754A-1	Sequence 1, Appl
4	58	100.0	604	14	US-10-079-754A-4	Sequence 4, Appl
5	51	87.9	585	14	US-10-079-754A-2	Sequence 2, Appl
6	34	58.6	525	14	US-10-079-754A-6	Sequence 6, Appl
7	27	46.6	869	14	US-10-079-754A-3	Sequence 3, Appl
8	11	19.0	96	14	US-10-079-754A-5	Sequence 5, Appl
C 9	9	15.5	7726	13	US-09-812-350-48	Sequence 48, Appl
C 10	8	13.8	181	13	US-09-918-995-8815	Sequence 8815, Ap
C 11	8	13.8	339	13	US-10-424-599-92870	Sequence 92870, A
C 12	8	13.8	390	10	US-09-918-995-8225	Sequence 8225, Ap
C 13	8	13.8	464	10	US-09-918-995-8118	Sequence 8118, Ap
C 14	8	13.8	808	17	US-10-437-963-56029	Sequence 56029, A
C 15	8	13.8	870	15	US-10-198-846-5963	Sequence 5963, Ap
C 16	8	13.8	1485	13	US-10-425-114-31277	Sequence 31277, A
C 17	8	13.8	1605	15	US-10-198-846-13051	Sequence 13051, A
C 18	8	13.8	12072	17	US-10-283-975A-255	Sequence 255, App
C 19	7	12.1	261	17	US-10-437-963-67969	Sequence 67969, A
C 20	7	12.1	320	9	US-09-962-832-138	Sequence 138, App
C 21	7	12.1	320	9	US-09-954-531-1010	Sequence 1010, Ap
C 22	7	12.1	342	17	US-10-437-963-39267	Sequence 39267, A
C 23	7	12.1	349	9	US-09-770-791-721	Sequence 721, App
C 24	7	12.1	349	13	US-10-282-122A-16688	Sequence 16688, A
C 25	7	12.1	357	17	US-10-437-963-31469	Sequence 91469, A
C 26	7	12.1	358	9	US-09-917-800A-522	Sequence 522, App
C 27	7	12.1	358	12	US-10-152-319A-585	Sequence 585, App
C 28	7	12.1	358	16	US-10-191-803-1105	Sequence 1105, Ap
C 29	7	12.1	364	10	US-09-873-367C-617	Sequence 617, App
C 30	7	12.1	369	13	US-10-424-599-40304	Sequence 40304, A
C 31	7	12.1	412	9	US-09-860-670-20	Sequence 20, Appl
C 32	7	12.1	412	16	US-10-227-646-20	Sequence 20, Appl
C 33	7	12.1	427	9	US-09-960-352-10246	Sequence 10246, A
C 34	7	12.1	454	13	US-10-424-599-114944	Sequence 114944, A
C 35	7	12.1	457	17	US-10-115-635-304	Sequence 304, App
C 36	7	12.1	485	10	US-09-918-995-32017	Sequence 32017, A
C 37	7	12.1	500	10	US-09-991-936-148	Sequence 148, App
C 38	7	12.1	509	13	US-10-424-599-36128	Sequence 36128, A
C 39	7	12.1	520	13	US-10-027-632-60602	Sequence 60602, A
C 40	7	12.1	520	13	US-10-027-632-60603	Sequence 60603, A
C 41	7	12.1	520	13	US-10-027-632-61489	Sequence 61489, A
C 42	7	12.1	520	13	US-10-027-632-61490	Sequence 61490, A
C 43	7	12.1	520	13	US-10-027-632-61688	Sequence 61688, A
C 44	7	12.1	520	13	US-10-027-632-61689	Sequence 61689, A
C 45	7	12.1	520	13	US-10-027-632-62287	Sequence 62287, A

## ALIGNMENTS

### RESULT 1

US-10-079-754A-15  
; Sequence 15, Application US/10079754A  
; Publication No. US20020164625A1  
; GENERAL INFORMATION:  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Grigor, Murray R.  
; APPLICANT: Molenaar, Adrian J.  
; APPLICANT: Davis, Stephen R.  
; TITLE OF INVENTION: Compositions Isolated from Bovine Mammary Gland and Methods for Their Use  
; FILE REFERENCE: 11000.1068  
; CURRENT APPLICATION NUMBER: US/10/079,754A  
; CURRENT FILING DATE: 2002-02-19  
; PRIOR APPLICATION NUMBER: US 09/699,146  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: US 60,162,701  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR APPLICATION NUMBER: US 09/644,190  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: US 60,150,330  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 15

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 15  
LENGTH: 267  
TYPE: DNA  
ORGANISM: Bovine  
US-10-079-754A-15

#### Alignment Scores:

Pred. No.: 9,01e-54 Length: 267  
Score: 58.00 Matches: 58  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-10-079-754A-10 (1-58) x US-10-079-754A-15 (1-267)

QY 1 MetLysllePheIlePheValPheIleMetAlaLeuIleLeuAlaMetIleArgAlaAsp 20  
DB 44 ATGAAGATCTTTATCTTTGCTTCATTATGGCTCTCATCTAGCCATGATTAGAGCTGAT 103  
QY 21 SerSerGluGluLysArgHisArgLysArgLysLysHisHisArgGlyTyrPheGlnGln 40  
DB 104 TCATCTGAAGAGAAACGTCACAGGAACCGGAAAAACATCATAGAGGATATTTCAACAA 163  
QY 41 TyrGlnProTyrGlnArgTyrProLeuAsnTyrProProAlaTyrProPhePro 58  
DB 164 TACCAGCCATATCAACGATATCCATAAATATCTCTCGGTATCCATTTCCT 217

#### RESULT 2

US-10-079-623-200

Sequence 200, Application US/10079623  
Publication No. US20020169302A1  
GENERAL INFORMATION:  
APPLICANT: Havukkala, Ilkka J.  
APPLICANT: Glenn, Matthew  
APPLICANT: Grigor, Murray R.  
APPLICANT: Molenaar, Adrian J.  
TITLE OF INVENTION: Compositions isolated from bovine  
FILE REFERENCE: 11000.1044c3  
CURRENT APPLICATION NUMBER: US/10/079,623  
CURRENT FILING DATE: 2002-02-19  
NUMBER OF SEQ ID NOS: 370  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 200  
LENGTH: 267  
TYPE: DNA  
ORGANISM: Bovine  
US-10-079-623-200

#### Alignment Scores:

Pred. No.: 9,01e-54 Length: 267  
Score: 58.00 Matches: 58  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-10-079-754A-10 (1-58) x US-10-079-623-200 (1-267)

QY 1 MetLysllePheIlePheValPheIleMetAlaLeuIleLeuAlaMetIleArgAlaAsp 20  
DB 44 ATGAAGATCTTTATCTTTGCTTCATTATGGCTCTCATCTAGCCATGATTAGAGCTGAT 103  
QY 21 SerSerGluGluLysArgHisArgLysArgLysLysHisHisArgGlyTyrPheGlnGln 40  
DB 104 TCATCTGAAGAGAAACGTCACAGGAACCGGAAAAACATCATAGAGGATATTTCAACAA 163  
QY 41 TyrGlnProTyrGlnArgTyrProLeuAsnTyrProProAlaTyrProPhePro 58  
DB 164 TACCAGCCATATCAACGATATCCATAAATATCTCTCGGTATCCATTTCCT 217

#### RESULT 3

US-10-079-754A-1

Sequence 1, Application US/10079754A  
Publication No. US20020164625A1  
GENERAL INFORMATION:  
APPLICANT: Glenn, Matthew  
APPLICANT: Grigor, Murray R.  
APPLICANT: Molenaar, Adrian J.  
APPLICANT: Davis, Stephen R.  
TITLE OF INVENTION: Compositions isolated from Bovine  
FILE REFERENCE: 11000.1068  
CURRENT APPLICATION NUMBER: US/10/079,754A  
CURRENT FILING DATE: 2002-02-19  
PRIOR APPLICATION NUMBER: US 09/699,146  
PRIOR FILING DATE: 2000-10-27  
PRIOR APPLICATION NUMBER: US 60,162,701  
PRIOR FILING DATE: 1999-10-29  
PRIOR APPLICATION NUMBER: US 09/644,190  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: US 60,150,330  
PRIOR FILING DATE: 1999-08-23  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 505  
TYPE: DNA  
ORGANISM: Bovine  
US-10-079-754A-1

#### Alignment Scores:

Pred. No.: 1,66e-53 Length: 505  
Score: 58.00 Matches: 58  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-10-079-754A-10 (1-58) x US-10-079-754A-1 (1-505)

QY 1 MetLysllePheIlePheValPheIleMetAlaLeuIleLeuAlaMetIleArgAlaAsp 20  
DB 44 ATGAAGATCTTTATCTTTGCTTCATTATGGCTCTCATCTAGCCATGATTAGAGCTGAT 103  
QY 21 SerSerGluGluLysArgHisArgLysArgLysLysHisHisArgGlyTyrPheGlnGln 40  
DB 104 TCATCTGAAGAGAAACGTCACAGGAACCGGAAAAACATCATAGAGGATATTTCAACAA 163  
QY 41 TyrGlnProTyrGlnArgTyrProLeuAsnTyrProProAlaTyrProPhePro 58  
DB 164 TACCAGCCATATCAACGATATCCATAAATATCTCTCGGTATCCATTTCCT 217

#### RESULT 4

US-10-079-754A-4

Sequence 4, Application US/10079754A  
Publication No. US20020164625A1  
GENERAL INFORMATION:  
APPLICANT: Glenn, Matthew  
APPLICANT: Grigor, Murray R.  
APPLICANT: Molenaar, Adrian J.  
APPLICANT: Davis, Stephen R.  
TITLE OF INVENTION: Compositions isolated from Bovine  
FILE REFERENCE: 11000.1068  
CURRENT APPLICATION NUMBER: US/10/079,754A  
CURRENT FILING DATE: 2002-02-19  
PRIOR APPLICATION NUMBER: US 09/699,146  
PRIOR FILING DATE: 2000-10-27  
PRIOR APPLICATION NUMBER: US 60,162,701  
PRIOR FILING DATE: 1999-10-29  
PRIOR APPLICATION NUMBER: US 09/644,190  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: US 60,150,330

```
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 604
; TYPE: DNA
; ORGANISM: Bovine
US-10-079-754A-4

Alignment Scores:
Pred. No.: 1.97e-53 Length: 604
Score: 58.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-079-754A-10 (1-58) x US-10-079-754A-4 (1-604)

Qy 1 MetLysilePheilePheValPheileMetAlaLeuileLeuAlaMetileArgAlaAasp 20
Db 154 ATGAAGATCTTATCTTGTCTTCATTCATTCCTCCTAGCCATGATTAGAGCTGAT 213
Qy 21 SerSerGluGluLysArgHisArgLysArgLysHisArgLysHisArgLysHisArgLysHis 40
Db 214 TCATCTGAAGAGAAACGTCACAGGAAACGGAAACACATCATAGAGATATTTTCAACAA 273
Qy 41 TyrGlnProTyrGlnArgTyrProLeuAsnTyrProAlaTyrProPhePro 58
Db 274 TACCAGCCATATCAACGATATCCACTAAATATCTCTCTCGGTATCCATTCCT 327

RESULT 5
US-10-079-754A-2
; Sequence 2, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Bovine
US-10-079-754A-5

Alignment Scores:
Pred. No.: 1.5e-27 Length: 525
Score: 34.00 Matches: 34
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 58.62% Indels: 0
DB: 14 Gaps: 0

US-10-079-754A-10 (1-58) x US-10-079-754A-6 (1-525)

Qy 1 MetLysilePheilePheValPheileMetAlaLeuileLeuAlaMetileArgAlaAasp 20
Db 43 ATGAAGATCTTATCTTGTCTTCATTCATTCCTCCTAGCCATGATTAGAGCTGAT 102
Qy 21 SerSerGluGluLysArgHisArgLysArgLysHisHisHis 34
Db 103 TCATCTGAAGAGAAACGTCACAGGAAACGGAAACACATCAT 144

RESULT 7
US-10-079-754A-3
; Sequence 3, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
```

```
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 604
; TYPE: DNA
; ORGANISM: Bovine
US-10-079-754A-4

Alignment Scores:
Pred. No.: 1.97e-53 Length: 604
Score: 58.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-079-754A-10 (1-58) x US-10-079-754A-4 (1-604)

Qy 1 MetLysilePheilePheValPheileMetAlaLeuileLeuAlaMetileArgAlaAasp 20
Db 154 ATGAAGATCTTATCTTGTCTTCATTCATTCCTCCTAGCCATGATTAGAGCTGAT 213
Qy 21 SerSerGluGluLysArgHisArgLysArgLysHisArgLysHisArgLysHisArgLysHis 40
Db 214 TCATCTGAAGAGAAACGTCACAGGAAACGGAAACACATCATAGAGATATTTTCAACAA 273
Qy 41 TyrGlnProTyrGlnArgTyrProLeuAsnTyrProAlaTyrProPhePro 58
Db 274 TACCAGCCATATCAACGATATCCACTAAATATCTCTCTCGGTATCCATTCCT 327

RESULT 5
US-10-079-754A-2
; Sequence 2, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Bovine
US-10-079-754A-2

Alignment Scores:
Pred. No.: 7.03e-46 Length: 585
Score: 51.00 Matches: 51
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 87.93% Indels: 0
DB: 14 Gaps: 0

US-10-079-754A-10 (1-58) x US-10-079-754A-2 (1-585)

Qy 8 PheileMetAlaLeuileLeuAlaMetileArgAlaAaspSerSerGluGluLysArgHis 27
```

```

; LOCATION: (1) - (181)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-8815

Alignment Scores:      6.88
Pred. No.:             8
Score:                  8.00
Percent Similarity:    100.00%
Best Local Similarity: 100.00%
Query Match:           13.79%
Indels:                0
Matches:               0
Mismatch:              0
Length:               181

```

DB: 10 Gaps: 0

US-10-079-754A-10 (1-58) x US-09-918-995-8815 (1-181)

QY 10 MetAlaLeuIleLeuAlaMetIle 17

DB 92 ATGGCTTTGATTGGCGATGATC 115

## RESULT 11

US-10-424-599-92870/c  
; Sequence 92870, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 92870  
; LENGTH: 339  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_54877C.1  
US-10-424-599-92870

Alignment Scores:  
Pred. No.: 12.5 Length: 339  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 13.79% Indels: 0  
DB: 13 Gaps: 0

US-10-079-754A-10 (1-58) x US-10-424-599-92870 (1-339)

QY 16 MetIleArgAlaAspSerSerGlu 23

DB 225 ATGATCCGCGCAGACTCGAGTGAG 202

## RESULT 12

US-09-918-995-8225  
; Sequence 8225, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8225  
; LENGTH: 390  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-918-995-8225

Alignment Scores:  
Pred. No.: 14.3 Length: 390  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 13.79% Indels: 0  
DB: 10 Gaps: 0

US-10-079-754A-10 (1-58) x US-09-918-995-8225 (1-390)

QY 10 MetAlaLeuIleLeuAlaMetIle 17

DB 137 ATGGCTTTGATTGGCGATGATC 160

## RESULT 13

US-09-918-995-8118  
; Sequence 8118, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8118  
; LENGTH: 464  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(464)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-8118

Alignment Scores:  
Pred. No.: 16.9 Length: 464  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 13.79% Indels: 0  
DB: 10 Gaps: 0

US-10-079-754A-10 (1-58) x US-09-918-995-8118 (1-464)

QY 10 MetAlaLeuIleLeuAlaMetIle 17

DB 254 ATGGCTTTGATTGGCGATGATC 277

## RESULT 14

US-10-437-963-56029/c  
; Sequence 56029, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 56029  
; LENGTH: 808  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_57980C.1  
US-10-437-963-56029

Alignment Scores:                      Length:                      808  
 Pred. No.:                      28.8                      Matches:                      8  
 Score:                      8.00                      Conservative:                      0  
 Percent Similarity:                      100.00%                      Mismatches:                      0  
 Best Local Similarity:                      100.00%                      Indels:                      0  
 Query Match:                      13.79%                      Gaps:                      0  
 DB:                      17

US-10-079-754A-10 (1-58) x US-10-437-963-56029 (1-808)

Qy    11 AlaLeuIleLeuAlaMetIleArg 18  
 Db    90 GCCTCATCTGGCTATGATCAGG 67

## RESULT 15

US-10-198-846-5963/c  
 ; Sequence 5963, Application US/10198846  
 ; Publication No. US20030099974A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lillie, James  
 ; APPLICANT: Xu, Yongyao  
 ; APPLICANT: Wang, Youzhen  
 ; APPLICANT: Steinmann, Kathleen  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
 ; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
 ; FILE REFERENCE: MRI-049  
 ; CURRENT APPLICATION NUMBER: US/10/198,846  
 ; CURRENT FILING DATE: 2002-07-18  
 ; PRIOR APPLICATION NUMBER: 60/306,220  
 ; PRIOR FILING DATE: 2001-07-18  
 ; NUMBER OF SEQ ID NOS: 14084  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 5963  
 ; LENGTH: 870  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc.feature  
 ; LOCATION: 26, 412, 436, 541, 543, 646, 649, 653, 711, 730, 742, 764,  
 ; LOCATION: 768, 779, 781, 785, 795, 800, 804, 809, 811, 821, 823, 829,  
 ; LOCATION: 830, 832, 839, 851, 853, 866, 870  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-10-198-846-5963

Alignment Scores:                      Length:                      870  
 Pred. No.:                      30.9                      Matches:                      8  
 Score:                      8.00                      Conservative:                      0  
 Percent Similarity:                      100.00%                      Mismatches:                      0  
 Best Local Similarity:                      100.00%                      Indels:                      0  
 Query Match:                      13.79%                      Gaps:                      0  
 DB:                      15

US-10-079-754A-10 (1-58) x US-10-198-846-5963 (1-870)

Qy    10 MetAlaLeuIleLeuAlaMetIle 17  
 Db    121 ATGGCTTTGATTTGGCGATGATC 98

Search completed: August 16, 2004, 02:01:13  
 Job time : 329 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: August 6, 2004, 17:10:01 ; Search time 22 Seconds  
(without alignments)

2834.735 Million cell updates/sec

Title: US-10-079-754A-4

Perfect score: 1029

Sequence: 1 gaagattttcagttctata.....tatcaagcataaaaaaaa 604

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US10079754/runat\_06082004\_180756\_1728/app\_query.fasta\_1.775  
-DB=Issued\_Patents\_AA -QWMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -LOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10079754 @CGN 1 1 27 runat\_06082004\_180756\_1728 -NCPU=6 -ICPU=3  
-NC\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pcp:\*
- 2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pcp:\*
- 3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pcp:\*
- 4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pcp:\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS.COMB.pcp:\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles.pcp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	75	7.2	370	1	US-08-415-818-12
C 2	75	7.2	370	2	US-08-894-236-12
C 3	75	7.2	370	5	PCT-US96-01444-12
C 4	73.5	7.1	1912	4	US-09-495-714C-2
C 5	73.5	7.1	1977	4	US-09-495-714C-4
C 6	72.5	7.0	1985	4	US-09-495-714C-6
C 7	70	6.8	233	4	US-09-489-847-176
C 8	69.5	6.8	275	4	US-09-252-991A-19818
C 9	68.5	6.6	273	4	US-08-936-165A-395
C 10	68	6.6	482	4	US-09-107-532A-5672
C 11	67.5	6.6	926	1	US-08-159-340A-2
C 12	66.5	6.5	161	2	US-08-286-819A-25

13	66.5	6.5	161	3	US-08-980-357-25	Sequence 25, Appl
14	66.5	6.4	317	4	US-09-393-634-60	Sequence 60, Appl
15	66	6.4	396	4	US-09-079-030-219	Sequence 219, App
16	66	6.4	1056	4	US-09-079-030-217	Sequence 217, App
17	66	6.4	4536	4	US-09-180-422B-27	Sequence 27, Appl
18	66	6.4	4536	4	US-09-079-030-1	Sequence 1, Appl
19	66	6.4	4563	4	US-09-108-006C-1	Sequence 1, Appl
20	65	6.3	247	4	US-09-134-000C-5222	Sequence 5222, Ap
21	65	6.3	371	1	US-08-415-818-6	Sequence 6, Appl
22	65	6.3	371	2	US-08-894-236-6	Sequence 4, Appl
23	65	6.3	371	2	US-08-919-624-4	Sequence 6, Appl
24	65	6.3	371	5	PCT-US96-01444-6	Sequence 5146, Ap
25	64.5	6.2	62	4	US-09-134-001C-5146	Sequence 1, Appl
26	64.5	6.3	1608	4	US-09-568-407-1	Sequence 1244, Ap
27	64	6.2	160	4	US-09-198-432A-1244	Sequence 8193, Ap
28	64	6.2	491	4	US-09-489-039A-8193	Sequence 356, App
29	63.5	6.2	115	3	US-08-513-974B-356	Sequence 26, Appl
30	63.5	6.2	117	3	US-09-073-297-26	Sequence 2, Appl
31	63.5	6.2	124	1	US-08-307-499-2	Sequence 2, Appl
32	63.5	6.2	124	3	US-09-299-268-2	Sequence 20066 A
33	63.5	6.1	268	4	US-09-252-991A-20066	Sequence 13, Appl
34	63.5	6.2	348	1	US-08-118-270-13	Sequence 13, Appl
35	63.5	6.2	348	5	PCT-US93-08528-13	Sequence 7127, Ap
36	63.5	6.2	526	4	US-09-107-532A-7127	Sequence 3464, Ap
37	63.5	6.2	2404	4	US-09-134-001C-3464	Sequence 6911, Ap
38	63	6.1	177	4	US-09-107-532A-5911	Sequence 8, Appl
39	63	6.1	182	1	US-07-962-569A-8	Sequence 2, Appl
40	63	6.1	182	1	US-08-308-883-2	Sequence 2, Appl
41	63	6.1	182	1	US-08-730-163-2	Sequence 2, Appl
42	63	6.1	182	3	US-08-256-799-2	Sequence 2, Appl
43	63	6.1	182	3	US-08-462-437-2	Sequence 2, Appl
44	63	6.1	182	3	US-08-462-437-31	Sequence 31, Appl
45	63	6.1	346	3	US-09-073-297-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-08-415-818-12  
; Sequence 12, Application US/08415818  
; Patent No. 5621079  
; GENERAL INFORMATION:  
; APPLICANT: Cascieri, Margaret A.  
; APPLICANT: Linemeyer, David L.  
; APPLICANT: MacNeil, Douglas J.  
; APPLICANT: Shiao, Lin-Lin  
; APPLICANT: Strader, Catherine D.  
; APPLICANT: Tan, Carina P.  
; APPLICANT: Weinberg, David H.  
; TITLE OF INVENTION: NEUROPEPTIDE Y RECEPTOR  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Mary A. Appollina  
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
; CITY: Rahway  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/415,818  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/383,746  
; FILING DATE: 03-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Appollina, Mary A.  
; REGISTRATION NUMBER: 34,087

REFERENCE/DOCKET NUMBER: 19390  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-594-3462  
TELEFAX: 908-594-4720

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 370 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-415-818-12

Alignment Scores:  
Pred. No.: 0.892 Length: 370  
Score: 75.00 Matches: 49  
Percent Similarity: 35.51% Conservative: 27  
Best Local Similarity: 22.90% Mismatches: 70  
Query Match: 7.23% Indels: 68  
DB: 1 Gaps: 13

US-10-079-754A-4 (1-604) x US-08-415-818-12 (1-370)

```

QY 564 ATGGTAAGTCAGAGAAAGAGAGTAGAA-----ACCAATTACATAGTTCAGAGACTA 511
Db 59 llellePheLysLysGlnArgLysAlaGlnAsnPheThrSerlleLeulleAlaAsnLeu 78
QY 510 TCA-----TTATTATGTTTTAAATATATTGATTTCCAAAGTATGTGTG 463
Db 79 SerLeuSerAspThrLeuValCysValMetCysIleHisPheThrIlelleTyrThrLeu 98
QY 462 TCGAACTAC-----TTGACTGATTTTTTTTCTTTCTTTGTT 430
Db 99 MetAspHisTrpIlePheGlyAspThrMetCysArgLeuThrSerTyrValGlnSerVal 118
QY 429 -----TTATTCAATTGAAAGATGTTTTTCAACACAGAGAGTAGGAAAATC 379
Db 119 SerlleSerValSerllePheSerLeuValPheThrAlaValGluArgTyrGlnLeulle 138
QY 378 ATT-----GTGAAAAATCTCTTAATCATGTCTCTGTAGTTACTAAGCAGCATT 331
Db 139 ValAsnProArgGlyTrpLysProSerValThrHisAlaTyr----- 152
QY 330 TTAAGGAATGATACCGCAGGAGGATAATTAGTGG--ATATCGTTGATATGCTGGTA 274
Db 153 -----TTPGlylleThrLeulleIleTrpLeuPhe 161
QY 273 TTGTTGAAATATCTCTATGATGTTTTTCCGTTTCTCTGTGAGCTTCTCTTCAGATGA 214
Db 162 SerLeuLeuLeuSerllePheProPheLeuSerTyrHisLeuThr----- 176
QY 213 ATCAGCTTATCATCGCTAGCATGAGCAGC---CATAATGAGACACAAGATAAGATCTT 157
Db 177 -----AspGluProPheHisAsnLeuSerLeuProThrAspLeu 189
QY 156 CATATT-----TGGTGGAGTCGAGTCATGAAAGATGAAATT 121
Db 190 TyrThrHisGlnValAlaCysValGluAsnTrp-----ProSerLysLysAspArgLeu 207
QY 120 CTGAGCTTCAGGAGATGTTCTTGTCTAGAAAACCTGCTGTGTTTAAATACCTTGAATCC 61
Db 208 LeuPheThrThrSerLeuPheLeuGlnTyrPheValProLeu-----Gly 223
QY 60 TTCATTTCTTTTGTAAATACATCAGTTATGAGATCTTATTA 19
Db 224 PheilleulleCys---TyrLeuLysIleValilleCysLeu 236

```

RESULT 2

US-08-894-236-12  
Sequence 12, Application US/08894236  
Patent No. 5939263

GENERAL INFORMATION:  
APPLICANT: Cascieri, Margaret A.

APPLICANT: Linemeyer, David L.  
APPLICANT: MacNeil, Douglas J.  
APPLICANT: Shiao, Lin-Lin  
APPLICANT: Strader, Catherine D.  
APPLICANT: Tan, Carina P.  
APPLICANT: Weinberg, David H.  
TITLE OF INVENTION: NEUROPEPTIDE Y RECEPTOR  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Mary A. Appollina  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/894/236  
FILING DATE:  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/383,746  
FILING DATE: 03-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/415,818  
FILING DATE: 03-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Appollina, Mary A.  
REGISTRATION NUMBER: 34,087  
REFERENCE/DOCKET NUMBER: 19390Y  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-594-3462  
TELEFAX: 908-594-4720

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 370 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-894-236-12

Alignment Scores:  
Pred. No.: 0.892 Length: 370  
Score: 75.00 Matches: 49  
Percent Similarity: 35.51% Conservative: 27  
Best Local Similarity: 22.90% Mismatches: 70  
Query Match: 7.23% Indels: 68  
DB: 2 Gaps: 13

```

US-10-079-754A-4 (1-604) x US-08-894-236-12 (1-370)
QY 564 ATGGTAAGTCAGAGAAAGAGAGTAGAA-----ACCAATTACATAGTTCAGAGACTA 511
Db 59 llellePheLysLysGlnArgLysAlaGlnAsnPheThrSerlleLeulleAlaAsnLeu 78
QY 510 TCA-----TTATTATGTTTTAAATATATTGATTTCCAAAGTATGTGTG 463
Db 79 SerLeuSerAspThrLeuValCysValMetCysIleHisPheThrIlelleTyrThrLeu 98
QY 462 TCGAACTAC-----TTGACTGATTTTTTTTCTTTCTTTGTT 430
Db 99 MetAspHisTrpIlePheGlyAspThrMetCysArgLeuThrSerTyrValGlnSerVal 118
QY 429 -----TTATTCAATTGAAAGATGTTTTTCAACACAGAGAGTAGGAAAATC 379
Db 119 SerlleSerValSerllePheSerLeuValPheThrAlaValGluArgTyrGlnLeulle 138
QY 378 ATT-----GTGAAAAATCTCTTAATCATGTCTCTGTAGTTACTAAGCAGCATT 331

```





**QY**

369 AATCTCTTAATCATGTCCTGTAGTTACTAAGCAGCATTTTAAAGAAATGGATACGAGG 310  
:::||||| ||| |||||:::|::|  
**Db**

641 SerLeuSerAsnLeuValAlaSerLeuLeuAsnSerMet-----LysSerIlealaSer 658

```

110 ProAspProCysMetIleProLeuLysTyrLeuGlnAlaTyrProAlaSer----- 126
      ||||| ||| ||||| |||
337 GCTTAGTAACACAGACATGATTAGAGAGATTTTTCACAAATGAT----- 381
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
127 -----LeuGlnAspGlnValArgGlnLeuIleAlaGlnAsnArgLeuGlyGluTyr 143
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
382 -----TTTTCCTAC 390
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
144 LeuGluArgArgTyrProGlyArgHisAspValGlnSerAspLysAlaLeuTyrAlaTyr 163
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
391 TCTTTCTGTGTGTTGAAACACCATCTTCAAATGAATAAACACAGAAAAAATC--- 447
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
164 ThrMetAsnLeuLysGlnGluHisLeuArgAsnAlaProGlyLeuAspLysValLeuTyr 183
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
448 -----AGTCAAGTAGTTGCACACACATACTTTGGAATC 480
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
184 AspAsnLysLeuAspValValGlnArgAlaLeuGlyLeu 196
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



QY 513 -----CTATCATTTATGTTTAAATATTGATATTGATTTCCAGTATGTTGTGTC 460  
 Db 187 SerllelethrPheMetPheArgPheGluTyrPheThrThrArgTyrThrTyrSer 206  
 QY 459 AACPACTGTAGTCAATTTTCTTTCTTTGTTTATTTCAATTTGAAAGATGGTTTCAACACA 400  
 Db 207 AsnLeullethrAsnSerSerLeuTyrLeullePhe-----PheSerAla 221  
 QY 399 ACAGAAAGAGTA-----GGAAATCATTTGTG-----AAAATCTCTCTAATCAT 355  
 Db 222 ThrLysGlylleAlaThrAlaSerLeulleMetGluAlaGlnLysPheLeuThrHisHis 241  
 QY 354 GTC-----CTGTAGTTACTAAGCAGCATTTTA----- 328  
 Db 242 LysAsnleleArgArgLeuLeuValSerSerPheLeuPheleleTyrAsnValAsnProPhe 261  
 QY 327 -----AGAAATGGATACGAGGAGGATTAATTTAGTGATATCGTTGATATGGCTG 277  
 Db 262 AsnValSerArgTyrTyrIleSerTyrValValleLeuPheMetLeullePhePheThr 281  
 QY 276 GTATTGTTGAAATATCTCTCATG-----ATGTTTCTTCCGTTT 238  
 Db 282 LysLyslleLysValAsnGlnMerLeulleLeulleLeuGlyMetPhePhelePhe 301  
 QY 237 CCTGTGAGCTTTCTCTCAGA 217  
 Db 302 ProLeuLeuAsnPhePheArg 308

## RESULT 11

US-08-159-340A-2  
 ; Sequence 2, Application US/08159340A  
 ; Patent No. 5565352  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hoechsttrasser, Mark  
 ; APPLICANT: Papa, Feroz  
 ; TITLE OF INVENTION: DRUBIQUINATING ENZYME: COMPOSITIONS  
 ; TITLE OF INVENTION: AND METHODS  
 ; NUMBER OF SEQUENCES: 44  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Arnold, White & Durkee  
 ; STREET: P.O. Box 4433  
 ; CITY: Houston  
 ; STATE: Texas  
 ; COUNTRY: USA  
 ; ZIP: 77210

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/159,340A  
 FILING DATE: 24-NOV-1993  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Highlander, Steven L.  
 REGISTRATION NUMBER: 37,642  
 REFERENCE/DOCKET NUMBER: ARCD:112/HYL  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (512) 418-3000  
 TELEFAX: (512) 474-7577  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 926 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

US-08-159-340A-2

Alignment Scores:  
 Pred. No.: 9.1 Length: 926

Score: 67.50 Matches: 48  
 Percent Similarity: 39.11% Conservative: 22  
 Best Local Similarity: 26.82% Mismatches: 51  
 Query Match: 6.56% Indels: 58  
 DB: 1 Gaps: 10

US-10-079-754A-4 (1-604) x US-08-159-340A-2 (1-926)

QY 82 CAGCAGTCTTTCTAGCAAAAGAACATCTCTGAAGCATCAGAATTC-----ATCTTT 132  
 Db 390 GlnGlnSerAspAsnAspHisValLeuLysArgSerSerPheLysLysLeuPhe 409  
 QY 133 CATGACTGCCTCCACCAAAATATGAG-----ATCTTTATCTTTCTCTTC 177  
 Db 410 SerAsnTyrThrSerProAsnProLysAsnSerAsnSerAsnLeuTyr-SerIleSerSe 429  
 QY 178 ATTATGGCTCTCATCTAGCCATGATTAGAGCTGATTCATCTGAAGAAACGTCACAGG 237  
 Db 429 lleuSerlleSerSerSerProSerProleuProleuHisSerProAspProValLysGl 449  
 QY 238 AAACGGA-----AAAAACATCATAGAGA 261  
 Db 449 yAsnSerPheArglleAsnTyrProGluThrProHisLeuTyrLysAsnSerGluThrAs 469  
 QY 262 TATTTTC-----AACCAATACCGCCATATCAACGATAT 294  
 Db 469 pheMetThrAsnGlnArgGluGlnLeuAsnHisAsnSerPheAlaHislleAlaProIl 489  
 QY 295 CCACTAAATATCTCTCTCGCTATCCATTTCTCTAAATGCTGCTAGTAATACAGAC 354  
 Db 489 eAsn-----ThrLysAlaIleThrSerProSerArgTh 500  
 QY 355 A-----TGATTAGAGAGATTTTTCACATGATTTTCTCT-----ACTCTTTCTGTTGT 402  
 Db 500 xAlaThrProLysLeuGlnArg-----PheProGlnThrIleSerMetAs 515  
 QY 403 GTTGAAACACCATCTTTCAAAATGAATAAAACAAAGAAAAAATCAGTCAAGTAGTTCGA 462  
 Db 515 nLeu-----AsnMetAsnSerAsnGlyHisSerSerAlaThrSerThrIleGl 531  
 QY 463 CAACACATACTTGGATCAATCAATATCAATATTTTAAACATAATATAGATCTC 517  
 Db 531 nProSerCysLeuSerLeuSer-----AsnAsnAspSerLeu 543

## RESULT 12

US-08-286-819A-25  
 ; Sequence 25, Application US/08286819A  
 ; Patent No. 5871910

GENERAL INFORMATION:  
 APPLICANT: ARTHUR, MICHEL

APPLICANT: DUKTA-MALEN, SYLVIE  
 APPLICANT: MOLINAS, CATHERINE

APPLICANT: COURVALIN, PATRICE

TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE  
 TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPETIDES, IN PARTICULAR

TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR  
 TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS

NUMBER OF SEQUENCES: 54  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
 ADDRESSEE: P.C.

STREET: 1755 S. Jefferson Davis Highway, Suite 400  
 CITY: Arlington

STATE: Virginia  
 COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM: disk  
 MEDIUM TYPE: Floppy

COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/286,819A









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: August 6, 2004, 17:14:11 ; Search time 48.5 Seconds

(without alignments)  
7812.969 Million cell updates/sec

Title: US-10-079-754A-4

Perfect score: 1029

Sequence: 1 GAGGATTTTCAGTCTCTATA.....TATCAAGCATAAAAAAAA 604

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Dgapop 6.0 , Delcxt 7.0	

Searched: 1291235 seqs, 313682936 residues

Total number of hits satisfying chosen parameters: 2582470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+, n2p.model -DEV-xlp  
-Q-/cgn2\_1/USPTO-spool\_p/US10079754/runat\_06082004\_180758\_1746/app\_query.fasta\_1.775  
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINWATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blotsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10079754@cgn1.1.13 @runat 06082004\_180758\_1746  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
-----					

1	316	30.7	58	13	US-10-079-754A-7	Sequence 7, Appli
2	316	30.7	58	13	US-10-079-754A-10	Sequence 10, Appli
3	315	30.6	58	13	US-10-079-754A-8	Sequence 8, Appli
4	214.5	20.8	59	13	US-10-079-754A-12	Sequence 12, Appli
5	169	16.4	70	13	US-10-079-754A-9	Sequence 9, Appli
6	112	10.9	21	13	US-10-079-754A-11	Sequence 11, Appli
7	110	10.7	62	12	US-10-210-172-168	Sequence 168, App
8	95	9.2	51	9	US-09-917-340-19	Sequence 19, Appli
9	94.5	9.2	78	10	US-09-992-600A-6	Sequence 6, Appli
10	94.5	9.2	78	10	US-09-924-340-6	Sequence 6, Appli
11	94.5	9.2	78	10	US-09-992-095B-6	Sequence 6, Appli
12	94.5	9.2	78	10	US-09-999-570-6	Sequence 6, Appli
13	94.5	9.2	78	14	US-10-000-489-6	Sequence 6, Appli
14	94.5	9.2	78	14	US-10-000-986-6	Sequence 6, Appli
15	94.5	9.2	78	14	US-10-154-678-6	Sequence 6, Appli
16	78	7.6	885	16	US-10-149-310-246	Sequence 246, App
17	77	7.5	1830	12	US-10-424-599-147146	Sequence 147146,
18	76.5	7.4	197	9	US-09-811-284-197	Sequence 197, App
19	75	7.2	423	9	US-09-870-759-103	Sequence 103, App
20	75	7.2	423	10	US-09-751-708A-103	Sequence 103, App
21	74.5	7.2	174	16	US-10-437-963-149428	Sequence 149428, A
22	73.5	7.1	240	12	US-10-282-122A-54218	Sequence 54218, A
23	73.5	7.1	246	16	US-10-389-566-1227	Sequence 1227, Ap
24	73.5	7.1	382	15	US-10-369-493-5291	Sequence 5291, Ap
25	73	7.0	373	13	US-10-114-893-204	Sequence 204, App
26	73	7.1	1680	16	US-10-437-963-192175	Sequence 192175,
27	72.5	7.0	833	16	US-10-437-963-130430	Sequence 130430,
28	72.5	7.0	1428	16	US-10-437-963-193953	Sequence 193953,
29	72.5	7.0	1504	16	US-10-437-963-192104	Sequence 192104,
30	72	6.9	806	12	US-10-282-122A-63502	Sequence 63502, A
31	71	6.9	1201	15	US-10-214-529-7	Sequence 7, Appli
32	70.5	6.9	1363	16	US-10-437-963-192041	Sequence 192041,
33	70	6.8	233	12	US-10-351-334-176	Sequence 176, App
34	70	6.8	276	12	US-10-424-599-154373	Sequence 154373,
35	70	6.8	294	12	US-10-424-599-156848	Sequence 156848,
36	69.5	6.8	309	10	US-09-510-332-131	Sequence 131, App
37	69.5	6.8	614	12	US-10-282-122A-53563	Sequence 53563, A
38	69	6.7	229	12	US-10-282-122A-71761	Sequence 71761, A
39	69	6.7	317	12	US-10-424-599-159483	Sequence 159483,
40	69	6.7	406	15	US-10-369-493-2232	Sequence 2232, Ap
41	69	6.7	454	12	US-10-282-122A-52793	Sequence 52793, A
42	69	6.7	1517	16	US-10-437-963-127400	Sequence 127400,
43	68.5	6.6	273	9	US-09-939-980-395	Sequence 395, App
44	68	6.6	259	16	US-10-437-963-148811	Sequence 148811,
45	68	6.6	369	12	US-10-335-977-7067	Sequence 7067, Ap

#### ALIGNMENTS

#### RESULT 1

US-10-079-754A-7  
; Sequence 7, Application US/10079754A  
; Publication No. US20020164625A1  
; GENERAL INFORMATION:  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Grigor, Murray R.  
; APPLICANT: Molenaar, Adrian J.  
; APPLICANT: Davis, Stephen R.  
; TITLE OF INVENTION: Compositions Isolated from Bovine Mammary Gland and Methods for Their Use  
; FILE REFERENCE: 11000.1068  
; CURRENT APPLICATION NUMBER: US/10/079,754A  
; CURRENT FILING DATE: 2002-02-19  
; PRIOR APPLICATION NUMBER: US 09/699,146  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: US 60,162,701  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR APPLICATION NUMBER: US 09/644,190  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: US 60,150,330  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 7
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-7

Alignment Scores:
Pred. No.: 4.13e-29 Length: 58
Score: 316.00 Matches: 58
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 30.71% Indels: 0
DB: 13 Gaps: 0

US-10-079-754A-4 (1-604) x US-10-079-754A-7 (1-58)

QY 154 ATGAAGATCTTTATCTTTGCTTCATTATGCTCTCATCTACCATGATTAGAGCTGAT 213
DB 1 MetLysIlePheIlePheValPheIleMetAlaLeuIleLeuAlaMetIleArgAlaAsp 20

QY 214 TCATCTGAAGAGAACGTCACAGGAAACGGAAAAACATCATAGAGGATATTTCAACAA 273
DB 21 SerSerGluGluLysArgHisArgLysArgLysLysHisArgGlyTyrPheGlnGln 40

QY 274 TACCAGCCATATCAACGATATCCACTAAATATTCCTCTCGGTATCCATTCTCT 327
DB 41 TyrGlnProTyrGlnArgTyrProLeuAsnTyrProProAlaTyrProPhePro 58

RESULT 2
US-10-079-754A-10
; Sequence 10, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; TITLE OF INVENTION: Mammary Gland and Methods for Their Use
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; PRIOR FILING DATE: 2002-02-19
; PRIOR FILING DATE: 1999-10-29
; PRIOR FILING DATE: US 60,162,701
; PRIOR FILING DATE: 2000-08-22
; PRIOR FILING DATE: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR FILING DATE: US 60,150,330
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-8

Alignment Scores:
Pred. No.: 5.46e-29 Length: 58
Score: 315.00 Matches: 57
Percent Similarity: 100.00% Conservatives: 1
Best Local Similarity: 98.28% Mismatches: 0
Query Match: 30.61% Indels: 0
DB: 13 Gaps: 0

US-10-079-754A-4 (1-604) x US-10-079-754A-8 (1-58)

QY 154 ATGAAGATCTTTATCTTTGCTTCATTATGCTCTCATCTACCATGATTAGAGCTGAT 213
DB 1 MetLysIlePheIlePheValPheIleMetAlaLeuIleLeuAlaMetIleArgAlaAsp 20

QY 214 TCATCTGAAGAGAACGTCACAGGAAACGGAAAAACATCATAGAGGATATTTCAACAA 273
DB 21 SerSerGluGluLysArgHisArgLysArgLysLysHisArgGlyTyrPheGlnGln 40

QY 274 TACCAGCCATATCAACGATATCCACTAAATATTCCTCTCGGTATCCATTCTCT 327
DB 41 TyrGlnProTyrGlnArgTyrProLeuAsnTyrProProAlaTyrProPhePro 58

RESULT 4
US-10-079-754A-12
; Sequence 12, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; TITLE OF INVENTION: Mammary Gland and Methods for Their Use
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; CURRENT FILING DATE: 2002-02-19
```

```

; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-12

```

```

Alignment Scores:
Pred. No.: 7,4e-17 Length: 59
Score: 214.50 Matches: 47
Percent Similarity: 78.33% Conservative: 0
Best Local Similarity: 78.33% Mismatches: 1
Query Match: 20.85% Indels: 12
DB: 13 Gaps: 1

```

US-10-079-754A-4 (1-604) x US-10-079-754A-12 (1-59)

```

QY 154 ATGAAGATCTTATCTTCTTCATTATGCTCTCATCTAGCCATGATTAGAGCTGAT 213
|
|
|
Db 1 MetLysilePheHePheHePheHeMetAlaLeuLeuAlaMetileArgAlaAsp 20
|
|
|
QY 214 TCATCTGAAGAGAAACGCTCAGAGAAACGGAAAAACATCAT 255
|
|
|
Db 21 SerSerGluGluLysArgHisArgLysArgLysHisHisValAspArgSerProGlu 40
|
|
|
QY 256 -----AGAGATATTTTCAACATATCAGCCATATCAGCATATCCAC 298
|
|
|
Db 41 PheLeuLeuileGlnGlu-AspilePheAsnAsnThrSerHisileAsnAspileHis 59
|
|
|

```

```

RESULT 5
US-10-079-754A-9
; Sequence 9, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-9

```

```

Alignment Scores:
Pred. No.: 2.37e-11 Length: 70
Score: 169.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 97.06% Mismatches: 0

```

```

Query Match: 16.42% Indels: 0
DB: 13 Gaps: 0

```

US-10-079-754A-4 (1-604) x US-10-079-754A-9 (1-70)

```

QY 154 ATGAAGATCTTATCTTCTTCATTATGCTCTCATCTAGCCATGATTAGAGCTGAT 213
|
|
|
Db 1 MetLysilePheHePheHePheHeMetAlaLeuLeuAlaMetileArgAlaAsp 20
|
|
|
QY 214 TCATCTGAAGAGAAACGCTCAGAGAAACGGAAAAACATCAT 255
|
|
|
Db 21 SerSerGluGluLysArgHisArgLysArgLysHisHis 34
|
|
|

```

#### RESULT 6

```

US-10-079-754A-11
; Sequence 11, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Bovine
US-10-079-754A-11

```

```

Alignment Scores:
Pred. No.: 0.000143 Length: 21
Score: 112.00 Matches: 18
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 94.74% Mismatches: 0
Query Match: 10.88% Indels: 0
DB: 13 Gaps: 0

```

US-10-079-754A-4 (1-604) x US-10-079-754A-11 (1-21)

```

QY 265 TTTCACATATCAGCCATATCAACGATATCCACTAAATTCCTCCGCGTATCCA 321
|
|
|
Db 1 PheGlnGlnTyxGlnProTyxGluArgTyxProLeuAsnTyxProAlaLayxPro 19
|
|
|

```

#### RESULT 7

```

US-10-210-172-168
; Sequence 168, Application US/10210172
; Publication No. US20040043928A1
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol
; APPLICANT: Rieger, Daniel
; APPLICANT: Shinkets, Richard
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Li, Li
; APPLICANT: Ji, Weizhen
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Casman, Stacie

```

APPLICANT: Voss, Edward  
APPLICANT: Boldog, Ferenc  
APPLICANT: Gorman, Linda  
APPLICANT: Leite, Mario  
APPLICANT: Vernet, Corine  
APPLICANT: Anderson, David  
APPLICANT: Guo, Xiaojia  
APPLICANT: Zhong, Mei  
APPLICANT: Gezlach, Valerie  
APPLICANT: Hjal, Tord  
APPLICANT: Rastelli, Luca  
APPLICANT: Sytek, Kimberly  
APPLICANT: Edinger, Shomit  
APPLICANT: Ellerman, Karen  
APPLICANT: Malyankar, Uriel  
APPLICANT: MacDougall, John  
APPLICANT: Stone, David  
APPLICANT: Alsbrook II, John  
APPLICANT: Leplev, Denise et al.  
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
FILE REFERENCE: 21402-416 A  
CURRENT APPLICATION NUMBER: US/10/210,172  
CURRENT FILING DATE: 2001-08-01  
PRIOR APPLICATION NUMBER: 60/309,501  
PRIOR FILING DATE: 2001-08-02  
PRIOR APPLICATION NUMBER: 60/323,994  
PRIOR FILING DATE: 2001-09-21  
PRIOR APPLICATION NUMBER: 60/373,814  
PRIOR FILING DATE: 2002-04-19  
PRIOR APPLICATION NUMBER: 60/310,291  
PRIOR FILING DATE: 2001-08-03  
PRIOR APPLICATION NUMBER: 60/310,951  
PRIOR FILING DATE: 2001-08-08  
PRIOR APPLICATION NUMBER: 60/310,544  
PRIOR FILING DATE: 2001-08-07  
PRIOR APPLICATION NUMBER: 60/311,292  
PRIOR FILING DATE: 2001-08-09  
PRIOR APPLICATION NUMBER: 60/311,979  
PRIOR FILING DATE: 2001-08-13  
PRIOR APPLICATION NUMBER: 60/313,201  
PRIOR FILING DATE: 2001-08-17  
PRIOR APPLICATION NUMBER: 60/312,892  
PRIOR FILING DATE: 2001-08-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 327  
SOFTWARE: Curaseq1 version 0.1  
SEQ ID NO 168  
LENGTH: 62  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-210-172-168  
Alignment Scores:  
Pred. No.: 0.000307 Length: 62  
Score: 110.00 Matches: 27  
Percent Similarity: 61.82% Conservative: 7  
Best Local Similarity: 49.09% Mismatches: 19  
Query Match: 10.69% Indels: 2  
Gaps: 2  
DB:  
US-10-079-754a-4 (1-604) x US-10-210-172-168 (1-62)  
QY 154 ATGAAGATCTTTAATCTTTGTTCTTATGCTCTCATCTAGCCATGATAGAGCTGAT 213  
Db 1 MetLysPheLeuValPheAlaPheLeuAlaLeuMetValSerMetIleGlyAlaasp 20  
QY 214 TCATCTGAAGAGAACGTCACAGGAAACCGAAAAACATCATAGAGGATATTTCAACA 273  
Db 21 SerSerGluGluLysPheLeuArgGileGlyArgPheGlyTyrGlyTyr---GlyPro 39  
QY 274 TACAGCCATATCAACGATATCCACTAAATATCTCTCTGGGTAT 318  
Db 40 TyrGlnProValProGluGlnProLeu---TyrProGlnProTyr 53

RESULT 8  
US-09-917-340-19  
Sequence 19, Application US/09917340  
Patent No. US20020090369A1  
GENERAL INFORMATION:  
APPLICANT: Murphy, Christopher J.  
APPLICANT: McAnulty, Jonathan F.  
APPLICANT: Reid, Ted W.  
TITLE OF INVENTION: Transplant Media  
FILE REFERENCE: TPLANT-08468  
CURRENT APPLICATION NUMBER: US/09/917,340  
CURRENT FILING DATE: 2001-07-29  
PRIOR APPLICATION NUMBER: 60/221,632  
PRIOR FILING DATE: 2000-07-28  
PRIOR APPLICATION NUMBER: 60/249,602  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/290,932  
PRIOR FILING DATE: 2001-05-15  
NUMBER OF SEQ ID NOS: 96  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 19  
LENGTH: 51  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-917-340-19  
Alignment Scores:  
Pred. No.: 0.0191 Length: 51  
Score: 95.00 Matches: 22  
Percent Similarity: 59.57% Conservative: 6  
Best Local Similarity: 46.81% Mismatches: 13  
Query Match: 9.23% Indels: 6  
Gaps: 2  
DB:  
US-10-079-754a-4 (1-604) x US-09-917-340-19 (1-51)  
QY 154 ATGAAGATCTTTAATCTTTGTTCTTATGCTCTCATCTAGCCATGATAGAGCTGAT 213  
Db 1 MetLysPhePheValPheAlaLeuAlaLeuMetLeuSerMetThrGlyAlaasp 20  
QY 214 TCATCTGAAGAGAACGTCAC---AGGAAACGGAAAAACAT-----CAT 255  
Db 21 SerHisAlaLysArgHisGlyTyrLysArgLysPheHisGluLysHisHisSerHis 40  
QY 256 AGAGGATATTTTCAACAATAC 276  
Db 41 ArgGlyTyrArgSerAsnTyr 47  
RESULT 9  
US-09-992-600A-6  
Sequence 6, Application US/09992600A  
Publication NO. US20030027161A1  
GENERAL INFORMATION:  
APPLICANT: Benjanin, Stephanie  
APPLICANT: Tanaka, Hiroaki  
TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
FILE REFERENCE: 91 US4.DIV  
CURRENT APPLICATION NUMBER: US/09/992,600A  
CURRENT FILING DATE: 2001-11-13  
PRIOR APPLICATION NUMBER: US 09/924,340  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: PCT/IB01/01715  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: US 60/305,456  
PRIOR FILING DATE: 2001-07-13  
PRIOR APPLICATION NUMBER: US 60/302,277  
PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: US 60/298,698  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 60/293,574  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 114





```
DB: 14 Gaps: 1
US-10-079-754A-4 (1-604) x US-10-000-986-6 (1-78)

Qy 154 ATGAAGATCTTTATCTTCTTCTTCATTCATTCCTGCTCTCATCTAGCCATGATTAGAGCTGAT 213
Db 1 MetLysPhePheValPheAlaLeuValLeuAlaLeuMetIleSerAlaAsp 20
Qy 214 TCATCTGAAGAGAAACGCTCAGGAAACGGGAAAAACATCATAGAGGATATTTTCAACAA 273
Db 21 SerHisGluLysArgHisGlyTyrArgArgLysPheHis----- 34
Qy 274 TACCAGCATATCAACGATATCCACTAAATTCCTCTCGGTAT 318
Db 35 ---GluLysHisHisSerTyrHisIleThrLeuLeuProLeuPhe 48

RESULT 15
US-10-154-678-6
; Sequence 6, Application US/10154678
; Publication No: US20030162186A1
; GENERAL INFORMATION:
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 182.US1.REG
; CURRENT APPLICATION NUMBER: US/10/154,678
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 6
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -19..-1
US-10-154-678-6

Alignment Scores:
Pred. No.: 0.0238 Length: 78
Score: 94.50 Matches: 18
Percent Similarity: 56.36% Conservative: 13
Best Local Similarity: 32.73% Mismatches: 17
Query Match: 9.18% Indels: 7
DB: 14 Gaps: 1

US-10-079-754A-4 (1-604) x US-10-154-678-6 (1-78)

Qy 154 ATGAAGATCTTTATCTTCTTCTTCATTCATTCCTGCTCTCATCTAGCCATGATTAGAGCTGAT 213
Db 1 MetLysPhePheValPheAlaLeuValLeuAlaLeuMetIleSerAlaAsp 20
Qy 214 TCATCTGAAGAGAAACGCTCAGGAAACGGGAAAAACATCATAGAGGATATTTTCAACAA 273
Db 21 SerHisGluLysArgHisGlyTyrArgArgLysPheHis----- 34
Qy 274 TACCAGCATATCAACGATATCCACTAAATTCCTCTCGGTAT 318
Db 35 ---GluLysHisHisSerTyrHisIleThrLeuLeuProLeuPhe 48
```

Search completed: August 6, 2004, 17:24:49  
Job time : 49.5 secs

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 14, 2004, 05:31:03 ; Search time 2769 Seconds  
(without alignments)  
9454.385 Million cell updates/sec

Title: US-10-079-754A-4  
Perfect score: 604  
Sequence: 1 gaagtatttcagttcrtata.....tatcaagcataaaaaaaa 604

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_scs.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pin.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sv.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	141	23.3	195534	2	AC134173	AC134173 Bos tauru
2	77	12.7	566	9	BC017835	BC017835 Homo sapi
3	77	12.7	325	9	AK131070	AK131070 Homo sapi
4	75.4	12.5	357	6	BD071675	BD071675 Secreted
5	70.4	11.7	438	6	AX588147	AX588147 Sequence
6	70.4	11.7	438	6	AX588483	AX588483 Sequence
7	70.4	11.7	438	6	AX616321	AX616321 Sequence
8	68.2	11.3	524	9	HUMHIS2X	M26665 Human hist
9	68.2	11.3	558	9	BC009791	BC009791 Homo sapi
10	68.2	11.3	1824	9	AK129614	AK129614 Homo sapi
11	68.2	11.3	2499	9	AK130503	AK130503 Homo sapi
12	68.2	11.3	2537	9	AK130505	AK130505 Homo sapi
13	67.6	11.2	542	9	HUMSTTRNA	M18371 Human stath
14	67.6	11.2	552	9	HUMSTATHA	M18078 Human stath
15	67.6	11.2	1584	6	AX747668	AX747668 Sequence
16	67.6	11.2	1584	9	AK092578	AK092578 Homo sapi
17	67.2	11.1	203	6	BD071749	BD071749 Secreted
18	66.6	11.0	480	9	HUMHIS1X	M26664 Human hist
19	65.6	10.9	486	4	AY154893	AY154893 Bos tauru
20	62.6	10.4	491	9	HUMHRPA	M18372 Human hist
21	60.8	10.1	516	6	BD071674	BD071674 Secreted
22	60.8	10.1	548	6	AX772840	AX772840 Sequence
23	60.4	10.0	231767	2	AC094531	AC094531 Rattus no
24	60.4	10.0	240446	2	AC131219	AC131219 Rattus no
C 25	60.4	10.0	284590	2	AC106950	AC106950 Rattus no
26	59.2	9.8	7550	9	HUMHIS102	L04132 Human hist
27	59.2	9.8	141568	9	AC063956	AC063956 Homo sapi
28	59.2	9.8	185969	2	AC069037	AC069037 Homo sapi
29	56.6	9.4	143842	2	AC134934	AC134934 Bos tauru
30	56.6	9.4	195334	2	AC134173	AC134173 Bos tauru
31	55.8	9.2	176	6	BD071751	BD071751 Secreted
C 32	55.8	9.2	250050	1	BX248584	BX248584 Blochmann
C 33	55.6	9.2	218675	2	AC015847	AC015847 Homo sapi
34	55	9.1	249943	3	AE014823	AE014823 Plasmodiu
C 35	54.6	9.0	49306	3	AC115606	AC115606 Dictyoste
36	54.4	9.0	253001	3	AE014834	AE014834 Plasmodiu
C 37	54.2	9.0	154071	3	AC115598	AC115598 Dictyoste
38	54.2	9.0	250823	3	AE014821	AE014821 Plasmodiu
C 39	53.6	8.9	13449	6	AX346287	AX346287 Sequence
C 40	53.6	8.9	250022	3	AE014824	AE014824 Plasmodiu
41	52.4	8.7	4500	3	AY392442	AY392442 Dictyoste
C 42	52.4	8.7	82139	3	AC115684	AC115684 Dictyoste
43	52	8.6	186934	5	BX004887	BX004887 Zebrafish
C 44	52	8.6	349980	6	AX344567	AX344567 Sequence
C 45	51.8	8.6	254436	3	AE014827	AE014827 Plasmodiu

ALIGNMENTS

RESULT 1  
AC134173/c  
LOCUS  
DEFINITION Bos taurus clone RP42-254113, WORKING DRAFT SEQUENCE, 26 unordered pieces.  
AC134173  
ACCESSION AC134173.1 GI:23306007  
VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
KEYWORDS Bos taurus (cow)  
SOURCE Bos taurus  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoides;  
Bovidae; Bovinae; Bos  
1 (bases 1 to 195534)  
REFERENCE  
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

Alabrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,  
Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,  
Bouck, J., Bowie, J., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,  
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
Carron, T.P., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, K., Chen, Z., Chowdhry, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
Gabis, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,  
Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,  
Lozato, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., Mathiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,  
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
Mosier, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuona, G.,  
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
Rives, M., Rojas, A., Rojubacan, I., Rolfe, M., Ruiz, S., Savery, G.,  
Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,  
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G. and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 195534)  
Worley, K.C.  
Direct Submission  
Submitted (24-SEP-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 195534)  
Worley, K.C.  
Direct Submission  
Submitted (27-SEP-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: VUAE  
Center clone name: RP42-254113  
----- Summary Statistics  
Sequencing vector: M13;  
Chemistry: Dye-primer Bodipy: 82% of reads  
Chemistry: Dye-terminator Big Dye: 18% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 185623 bases at least Q40  
Consensus quality: 190257 bases at least Q30  
Consensus quality: 192466 bases at least Q20  
Estimated insert size: 204493; sum-of-contigs estimation  
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation  
-----

\* NOTE: Estimated insert size may differ from sequence length

(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a "working draft" sequence. It currently  
consists of 26 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 2358: contig of 2358 bp in length  
2359 2458: gap of unknown length  
2459 5619: contig of 3161 bp in length  
5620 5719: gap of unknown length  
5720 8310: contig of 2591 bp in length  
8311 8410: gap of unknown length  
8411 11002: contig of 2592 bp in length  
11003 11002: gap of unknown length  
11003 14386: contig of 3284 bp in length  
14387 14886: gap of unknown length  
14887 17149: contig of 2663 bp in length  
17150 17249: gap of unknown length  
17250 21414: contig of 4165 bp in length  
21415 21514: gap of unknown length  
21515 22571: contig of 3657 bp in length  
22572 23271: gap of unknown length  
23272 30645: contig of 5374 bp in length  
30646 30745: gap of unknown length  
30746 36182: contig of 5437 bp in length  
36183 36282: gap of unknown length  
36283 41956: contig of 5674 bp in length  
41957 42056: gap of unknown length  
42057 48094: contig of 6038 bp in length  
48095 48194: gap of unknown length  
48195 53214: contig of 5020 bp in length  
53215 53314: gap of unknown length  
53315 5849: contig of 5535 bp in length  
5849 58949: gap of unknown length  
58950 66144: contig of 7195 bp in length  
66145 66244: gap of unknown length  
66245 73033: contig of 6789 bp in length  
73034 73133: gap of unknown length  
73134 80030: contig of 6897 bp in length  
80031 80130: gap of unknown length  
80131 87761: contig of 7831 bp in length  
87762 87861: gap of unknown length  
87862 95726: contig of 7865 bp in length  
95727 95826: gap of unknown length  
95827 105054: contig of 9228 bp in length  
105055 105154: gap of unknown length  
105155 116716: contig of 11562 bp in length  
116717 116816: gap of unknown length  
116817 128293: contig of 11477 bp in length  
128294 128393: gap of unknown length  
128394 138140: contig of 9747 bp in length  
138141 138240: gap of unknown length  
138241 150687: contig of 12447 bp in length  
150688 150787: gap of unknown length  
150788 166414: contig of 15627 bp in length  
166415 195534: gap of unknown length  
195535 195534: contig of 29020 bp in length.

FEATURES  
Source

1. 195534  
/organism="Bos taurus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9913"  
/clone="RP42-254113"

ORIGIN

Query Match 23.3%; Score 141; DB 2; Length 195534;  
Best Local Similarity 100.0%; Pred. No. 2.3e-17;  
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 GAAGTATTTCGTTCTATATATAGATCTCTACTGATGTAATACAAACAAATGAA 60

FEATURES	source	Location/Qualifiers
	1..566	Location/Qualifiers
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="MGC:22502 IMAGE:4289874"	
	/tissue_type="Skeletal Muscle"	
	/clone_lib="NIH MGC_81"	
	/lab_host="DH10B"	
	/notes="vector: pDNR-LIB"	
gene	1..566	
	/gene="HTN1"	
	/notes="synonym: HIS1"	
	/db_xref="LocusID:3346"	
	/db_xref="MIM:142701"	
CDS	68..241	
	/codon_start=1	
	/product="histatin 1"	
	/protein_id="AAH17835.1"	
	/db_xref="GI:17389615"	
	/db_xref="LocusID:3346"	
	/translation="MKFFVFLVLALMISMISADSHKRIHGYRRKFHKHSHREPP FYGDYGSNLYDN"	
ORIGIN		
Query Match	12.7%	Score 77; DB 9; Length 566;
Best Local Similarity	59.6%	Pred. No. 7.3e-05;
Matches	150; Conservative	0; Mismatches 95; Indels 6; Gaps 1;
QY	120	GAATTTCATCTTTCATGACTGGACCTCCACCAATATGAAGATCTTTATCTTTGCTTCAT 179
Db	34	GACTCTCTCTTCAGTAAAGGACTCAGCACTATGAAGTTTTTTGCTTGTCTTAGT 93
QY	180	TATGGCTTCTCATCTAGCCATGATTAGACTGATTCATCTGAAGAGAAAACGTCACAGAA 239
Db	94	CTTGCTCTCATGATTTCCATGATTAGCGCTGATTCACATGAAGAAGACATCATGGGTA 153
QY	240	ACGGA AAAACATCATAGAGGATATTTTCAACAATACAGCCATATCAACGATATCCACT 299
Db	154	TAGAAGAAAATTCAT-----GAAAAGCATCATTCACATCGAGAATTTCCATTTATGG 207
QY	300	AAATTATCTCTCGGTATCCATTTCTTAAATGCTGTTAGTAACATACAGGACATGAT 359
Db	208	GGACTATGGATCAATATCTATGACAAATGATATCCTTAGTAGTAATCATGGGCATGAT 267
QY	360	TAGAGAGATTT 370
Db	268	TATAGAGTTT 278
RESULT 3		
AK131070		
LOCUS	AK131070	3235 bp mRNA linear PRI 10-SEP-2003
DEFINITION	Homo sapiens cDNA FLJ29036 fis, clone SLV00594.	
ACCESSION	AK131070	
VERSION	AK131070.1	GI:34528324
KEYWORDS	cligo capping; fis (full insert sequence).	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
	1 Kanehori,K., Iehibashi,T., Chiba,Y., Fujimori,K., Hiraoka,S.,	
	Tanai,H., Watanabe,S., Ishida,S., Ono,Y., Horuta,T., Watanabe,M.,	
	Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S., Morinaga,M.,	
	Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,	
	Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isogai,T. and	
	Sugano,S.	
	2 Sugano,S. and Suzuki,Y.	
TITLE	NEDO human cDNA sequencing project	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 3235)	
AUTHORS	Sugano,S. and Suzuki,Y.	
TITLE	Direct Submission	

JOURNAL Submitted (31-JUL-2003) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan (E-mail: flicdn@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.

FEATURES  
source  
1..3235  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="SLV00594"  
/tissue type="salivary gland"  
/clone.lib="SLV"  
/note="cloning vector: pME18SPL3"

## ORIGIN

Query Match 12.7%; Score 77; DB 9; Length 3235;  
Best Local Similarity 59.8%; Pred.No. 5.5e-05;  
Matches 150; Conservative 0; Mismatches 95; Indels 6; Gaps 1;  
QY 120 GAATTTTCATCTTTTCATGACTGGACTCCACCAATATGAGATCTTTATCTTTGCTTCAT 179  
Db GACTCTCTCTTGAGTAAAGGACTGAGCACTATGAGATTTTTCCTTTCTTTAGT 2789  
QY 180 TATGGCTCTCATCTACCCATGATAGAGCTGATTCATCTGAAGAGAAAGCTCACAGGAA 239  
Db CTTCGCTCTCATGATTTCCATGATTAGCGCTGATTCACATGAAAGAGACATCATGGTA 2849  
QY 240 ACGGAAAACATCATAGAGGATTTTCAACATACCCCATATCAAGCATATCCACT 299  
Db TAGAAGAAATTCAT-----GAAAGCATATTCATCGAGATTTCCATTTATGG 2903  
QY 300 AAATTATCTCTCGTATCCATTTCTTAAATGCTGTAGTAACACAGGACATGAT 359  
Db GGACTATGGATCAATATCTATATGACATTCGATATCTTATGATTCATGAGGCGCATGAT 2963  
QY 360 TAGAGAGATTT 370  
Db 2964 TATAGAGGTTT 2974

RESULT 4  
LOCUS BD071675 357 bp DNA linear PAT 27-AUG-2002  
DEFINITION Secreted expressed sequence tags (seSTs).  
ACCESSION BD071675  
VERSION BD071675.1 GI:22617278  
KEYWORDS JP 2001519667-A/485.  
SOURCE unidentified  
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 357)  
AUTHORS Jacobs,K., Mccoy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D., Treacy,M., Spaulding,V. and Agostino,M.J.  
TITLE Secreted expressed sequence tags (seSTs).  
JOURNAL Patent: JP 2001519667-A 485 23-OCT-2001;  
GENETICS INSTITUTE INC  
COMMENT OS Unidentified  
PN JP 2001519667-A/485  
PD 23-OCT-2001  
PF 10-APR-1998 JP 1998543069  
PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI DAVID MERBERG.  
PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC  
C12N15/12, C12N25/10, C07K14/47, C12Q1/68, A61K38/17 CC Strandedness:

Double:  
CC Topology: Linear;  
CC Secreted expressed sequence tags (seSTs)  
FH Key location/Qualifiers  
FT source 1..357  
FT /organism="Unidentified".  
Location/Qualifiers  
1..357  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

## FEATURES

source  
1..357  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
84..320  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD61372.1"  
/db\_xref="GI:27899823"  
/translation="MKFFVFLVIALMISMISADSEKHHGYRRKPEKHHSHYITL  
LPFEESKSNANEKHNYLLYTLCFRLAPSVIT"  
sig\_peptide 84..140  
/note="Von Heijne matrix score 7.64030745849671 seq  
ALVIALMISMISA/DS"  
polya\_signal 397..402  
polya\_site 421..438

## ORIGIN

Query Match 12.5%; Score 75.4; DB 6; Length 357;  
Best Local Similarity 59.4%; Pred.No. 0.00016;  
Matches 149; Conservative 0; Mismatches 96; Indels 6; Gaps 1;  
QY 120 GAATTTTCATCTTTTCATGACTGGACTCCACCAATATGAGATCTTTATCTTTGCTTCAT 179  
Db GACTCTCTCTTGAGTAAAGGACTCAGCCAACTATGAGATTTTTCCTTTCTTTAGT 107  
QY 180 TATGGCTCTCATCTACCCATGATAGAGCTGATTCATCTGAAGAGAAAGCTCACAGGAA 239  
Db CTTCGCTCTCATGATTTCCATGATTAGCGCTGATTCACATGAAAGAGACATCATGGTA 167  
QY 240 ACGGAAAACATCATAGAGGATTTTCAACATACAGCCATATCAAGATATCCACT 299  
Db TAGAAGAAATTCAT-----GAAAGCATATTCATCGAGATTTCCATTTATGG 221  
QY 300 AAATTATCTCTCGTATCCATTTCTTAAATGCTGTAGTAACACAGGACATGAT 359  
Db GGACTTGGATCAATATCTATATGACATTCGATATCTTATGATTCATGAGGCGCATGAT 281  
QY 360 TAGAGAGATTT 370  
Db 282 TATAGAGGTTT 292

## FEATURES

source  
1..438  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
84..320  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD61372.1"  
/db\_xref="GI:27899823"  
/translation="MKFFVFLVIALMISMISADSEKHHGYRRKPEKHHSHYITL  
LPFEESKSNANEKHNYLLYTLCFRLAPSVIT"  
sig\_peptide 84..140  
/note="Von Heijne matrix score 7.64030745849671 seq  
ALVIALMISMISA/DS"  
polya\_signal 397..402  
polya\_site 421..438

LOCUS	AX616321	438 bp	DNA	linear	PAT 20-FEB-2000
DEFINITION	Sequence 5 from Patent WO02094864.				
ACCESSION	AX616321				
VERSION	AX616321.1	GI:28447365			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
TITLE	Bejanin, S. and Tanaka, H.				
JOURNAL	Human cdnas and proteins and uses thereof				
FEATURES	Patent: WO 02094864-A 5 28-NOV-2002;				
source	GENSET (FR)				
5'UTR	Location/Qualifiers				
CDS	1..438				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
	1..83				
	84..320				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD67625.1"				
	/db_xref="GI:28447366"				
	/db_xref="REMBL:CAD67625"				
	/translation="MKFFVFALVLAIMISADSHKHHGRRKFKHKKHSHYHLL"				
	LPLFESSKSNAMKHYLLYLTCRILAFSIVT"				
	3'UTR	318..438			
	polyA_signal	397..402			
	polyA_site	423..438			
ORIGIN					
	Query Match	11.7%;	Score 70.4;	DB 6;	Length 438;
	Best Local Similarity	69.9%;	Pred. No. 0.0013;		
	Matches	95;	Conservative	0;	Mismatches 41;
					Indels 0;
					Gaps 0;
QY	120	GAATTCATCTTTCATGACTGGACTCCACCAATATGAGATCTTTATCTTTGCTTCAT	179		
Db	50	GACTCTCCTCTTGAGTAAAGGACTCAGCACTATGAGTTTTTTGCTTCTTTAGT	109		
QY	180	TATGGCTCTCATCTAGCCATGATTAGAGCTGATTCATCTGAAGAGAAAGCTCACAGAA	239		
Db	110	CTTGGCTCTCATGATTTCCATGATTAGGCTGATTCACATGAAAGAGACATCATGGTA	169		
QY	240	ACGGAAAAACATCAT	255		
Db	170	TAGAAGAAAAATTCAT	185		
RESULT 8					
HUMHIS2X					
LOCUS	Human histatin 2 (HIS2)	524 bp	mRNA	linear	PRI 31-DEC-1994
DEFINITION	Human histatin 2 (HIS2) mRNA, complete cds.				
ACCESSION	M26665				
VERSION	M26665.1	GI:292145			
KEYWORDS	histatin 2.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
TITLE	1 (bases 1 to 524)				
JOURNAL	Histatins, a family of salivary histidine-rich proteins, are				
EDITION	encoded by at least two loci (HIS1 and HIS2)				
PUBMED	Biochem. Biophys. Res. Commun. 160 (2), 495-502 (1989)				
COMMENT	Original				
FEATURES	source text: Homo sapiens parotid gland cdna to mRNA.				
source	Location/Qualifiers				
	1..524				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				

```

/db_xref="taxon:9606"
/tissue_type="parotid gland"
1. .524
/gene="HIS2"
72. .227
/gene="HIS2"
/codon_start=1
/product="histatin 2"
/protein_id="AA58646.1"
/db_xref="GI:292146"
/translation="MKFFVFLALMLMTGADSHAKRHGKRFHEKHHSHRGYR
SNLYDN"
polya_signal
506. .511
/gene="HIS2"

ORIGIN
Query Match 11.3%; Score 68.2; DB 9; Length 524;
Best Local Similarity 65.4%; Pred. No. 0.0034;
Matches 100; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 120 GAATTTTCATCTTTTCATGACGCTCCACCAATATGAAGATCTTTATCTTTGCTTTTCAT 179
Db 38 GATTCCTCTTTGAGTAAAGGACTGACCACTATGAAGTTTTTTTGTTCCTTAAT 97

Qy 180 TATGGCTCTCATCTCCATGATTTAGAGTGATTCATCTGAAGAAACGTCACAGGAA 239
Db 98 CTGGCTCTCATGCTTTCCATGACTGGAGCTGATTCACATGCAAGAGACATCATGGGTA 157

Qy 240 ACGGAAACATCATAGAGGATTTTCAACA 272
Db 158 TAAAGAAATTCATGAAAGCATCTTCA 190

RESULT 9
BC009791 Homo sapiens histatin 3, mRNA (CDNA clone MGC:13578 IMAGE:4293405),
LOCUS complete cds.
ACCESSION BC009791
VERSION BC009791.1 GI:14602560
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 558)
Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Scapleton M., Soares M.B., Bonaldo M.P., Casavant T.L.,
Schuetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S.,
Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J.,
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Heintz E., Kettelman M., Madan A., Rodriguez S.,
Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y.,
Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield V.S., Krzywicki M.I., Skalska U., Smallos D.E.,
Schnerch A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 558)
Straussberg R.
AUTHORS Direct Submission
TITLE Submitted (02-JUL-2003) National Institutes of Health, Mammalian
JOURNAL

```

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)

Tissue Procurement: CLONTECH

CDNA Library Preparation: CLONTECH Laboratories, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbiology.org>

Contact: [amadansystemsbiology.org](mailto:amadansystemsbiology.org)  
Anup Madan, Jessica Fahey, Erin Heiton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IMAGE Plate: 19 Row: 0 Column: 1

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4557652.

Location/Qualifiers

1. .558  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:13578 IMAGE:4293405"  
/tissue\_type="Skeletal Muscle"  
/clone\_lib="NIH MGC\_81"  
/lab\_host="DH10B"  
/notf="Vector: pDNR-LIB"  
1. .558  
/gene="HIS2"  
/note="synonyms: HIS2, HTN2, HTN5"  
/db\_xref="LocusID:3347"  
/db\_xref="MIM:142702"  
76. .231  
/codon\_start=1  
/product="histatin 3"  
/protein\_id="AAH09791.1"  
/db\_xref="GI:14602561"  
/db\_xref="LocusID:3347"  
/translation="MKFFVFLALMLMTGADSHAKRHGKRFHEKHHSHRGYR  
SNLYDN"

Query Match 11.3%; Score 68.2; DB 9; Length 558;  
Best Local Similarity 65.4%; Pred. No. 0.0034;  
Matches 100; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 120 GAATTTTCATCTTTTCATGACGCTCCACCAATATGAAGATCTTTATCTTTGCTTTTCAT 179  
Db 42 GATTCCTCTTTGAGTAAAGGACTGACCACTATGAAGTTTTTTTGTTCCTTAAT 101

Qy 180 TATGGCTCTCATCTCCATGATTTAGAGTGATTCATCTGAAGAAACGTCACAGGAA 239  
Db 102 CTGGCTCTCATGCTTTCCATGACTGGAGCTGATTCACATGCAAGAGACATCATGGGTA 161

Qy 240 ACGGAAACATCATAGAGGATTTTCAACA 272  
Db 162 TAAAGAAATTCATGAAAGCATCTTCA 194

RESULT 10  
AK129614  
LOCUS AK129614  
DEFINITION AK129614  
ACCESSION AK129614  
VERSION AK129614.1 GI:34526191  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1  
**REFERENCE**  
**AUTHORS**  
 Kanehori,K., Ishibashi,T., Chiba,Y., Fujimori,K., Hiraoka,S., Tanai,H., Watanabe,S., Ishida,S., Ono,Y., Hotuta,T., Watanabe,M., Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S., Morinaga,M., Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isogai,T. and Sugano,S.  
**TITLE**  
 NEDO human cDNA sequencing project  
**JOURNAL**  
 Unpublished  
**REFERENCE**  
 2 (bases 1 to 1824)  
**AUTHORS**  
 Sugano,S. and Suzuki,Y.  
**TITLE**  
 Direct Submission  
**JOURNAL**  
 Submitted (31-JUL-2003) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan  
 (E-mail:flcnao@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)  
**COMMENT**  
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing; Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.  
**FEATURES**  
 Location/Qualifiers  
 1..1824  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="SLV07780"  
 /tissue\_type="salivary gland"  
 /clone\_lib="SLV"  
 /note="cloning vector: pME18SFL3"  
**ORIGIN**  
 Query Match 11.3%; Score 68.2; DB 9; Length 1824;  
 Best Local Similarity 65.4%; Pred. No. 0.0028;  
 Matches 100; Conservative 0; Mismatches 53; Indels 0; Gaps 0;  
 QY 120 GAATTTTCATCTTTCATGACTGCTGACTCCACCAAAATGACGATCTTTATCTTTGCTTCAT 179  
 DB 1338 GATTCTCTCTTTGAGTAAGGATCTAGCCAACTATGAGTTTTCGTTTTCGTTTAAAT 1397  
 QY 180 TATGGCTCTCATCCTTAGCCATGATTAGAGCTGATTCATCTCTGAAGAGAAACGTCACAGAA 239  
 DB 1398 CTTCGCTCTCATGCTTTCCATGACTGGAGCTGATTCACATGTCACCAAGAGACATCATGGGTA 1457  
 QY 240 ACGGAAAAACATCATAGAGGATATTTTCACA 272  
 DB 1458 TAAAGGAAATTCCTATGAAGATCATTCACA 1490  
**RESULT 11**  
 AKI30503 2499 bp mRNA linear PRI 10-SEP-2003  
**LOCUS**  
**DEFINITION**  
 Homo sapiens cDNA FLJ26993 fis, clone SLV03847, highly similar to Histatin 3 precursor.  
**ACCESSION**  
 AKI30503  
**VERSION**  
 AKI30503.1 GI:34527320  
**KEYWORDS**  
 oligo capping; fis (full insert sequence).  
**SOURCE**  
 Homo sapiens (human)  
**ORGANISM**  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE**  
 1  
**AUTHORS**  
 Kanehori,K., Ishibashi,T., Chiba,Y., Fujimori,K., Hiraoka,S., Tanai,H., Watanabe,S., Ishida,S., Ono,Y., Hotuta,T., Watanabe,M., Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S., Morinaga,M., Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isogai,T. and Sugano,S.  
**TITLE**  
 NEDO human cDNA sequencing project  
**JOURNAL**  
 Unpublished

2 (bases 1 to 2499)  
Sugano,S. and Suzuki,Y.  
Direct Submission  
Submitted (31-JUL-2003) Sumio Sugano, Institute of Medical Science,  
University of Tokyo, Laboratory of Genome Structure, Human Genome  
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan  
(E-mail: flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,  
Fax:81-3-5449-5416)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction and 5'-end one pass sequencing; Institute of Medical  
Science, University of Tokyo, Laboratory of Genome Structure, Human  
Genome Center; 3'-end one pass sequencing: RAB; clone selection for  
full insert sequencing: RAB and Helix Research Institute.  
Location/Qualifiers  
1. .2499  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="SLV03847"  
/tissue\_type="salivary gland"  
/clone\_lib="SLV"  
/notes="cloning vector: pME18SFL3"

ORIGIN  
Query Match 11.3%; Score 68.2; DB 9; Length 2499;  
Best Local Similarity 55.4%; Pred. No. 0.0027;  
Matches 100; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 120 GAATTTTCATCTTTTCATCGACTGCGACTCCACAAATATGAAGATCTTTATCTTTGCTTTTCAT 179  
Db 2013 GATTTCCTCTTCAGTAAAGGACTCAGGCACTATGCAAGATTTTGTGTTTTCGCTTTAAT 2072

QY 180 TATGGCTCTCATCTTAGCCATGATTAGAGCTGATTCATCTGAAGAGAAAGCTCACAGGAA 239  
Db 2073 CTTGGCTCTCATGCTTTCCATGACTGAGCTGATTCACATGCGAAGAGACATCATGGGA 2132

QY 240 ACGGAAAAACATCATAGAGCATATTTTCAACA 272  
Db 2133 TAAAGAAATTCATGAAGGATCATTCACA 2165

RESULT 12  
AK130505  
LOCUS Homo sapiens cDNA FLJ26995 fis, clone SLV04017, highly similar to  
DEFINITION Homo sapiens histatin 3 (HTN3).  
ACCESSION AK130505.1 GI:34527322  
VERSION AK130505.1  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Kanehori,K., Ishibashi,T., Chiba,Y., Fujimori,K., Hiraoka,S.,  
Tanai,H., Watanabe,S., Ishida,S., Ono,Y., Hotuta,T., Watanabe,M.,  
Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S., Morinaga,M.,  
Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,  
Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isegai,T. and  
Sugano,S.  
NEDO human cDNA sequencing project  
JOURNAL UNpublished  
REFERENCE 2 (bases 1 to 2537)  
AUTHORS Sugano,S. and Suzuki,Y.  
TITLE Direct Submission  
JOURNAL Submitted (31-JUL-2003) Sumio Sugano, Institute of Medical Science,  
University of Tokyo, Laboratory of Genome Structure, Human Genome  
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan  
(E-mail: flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,  
Fax:81-3-5449-5416)  
NEDO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.

## FEATURES

source  
1..2537  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="SLV04017"  
/tissue\_type="salivary gland"  
/clone\_lib="SLV"  
/note="Cloning vector: pME18SFL3"

## ORIGIN

Query Match 11.3%; Score 68.2; DB 9; Length 2537;  
Best Local Similarity 65.4%; Pred. No. 0.0026;  
Matches 100; Conservative 0; Mismatches 53; Indels 0; Gaps 0;  
QY 120 GAATTCATCTTCATGACTGCTCCACCAAAATATGAAGATCTTTATCTTGTCTTCAT 179  
DB 2051 GATTCCTCTTGTAGTAAAGACTCAGCCAACTATGAAATTTTGTCTTTAT 2110  
QY 180 TATGGCTCTCATCTCCATGATTTAGAGCTGATTCATCTGAAGAGAAACGTCAAGGAA 239  
DB 2111 CTGGCTCTCATGTTTCCATGCTGAGCTGATTCATGCAAGAGACATCATGGTA 2170  
QY 240 ACGGAAACATCATGAGGATATTTCAACA 272  
DB 2171 TAAAGAAATTCATGAAAGCATCATTCACA 2203

## RESULT 13

HUMSTTRNA HUMSTTRNA 542 bp mRNA linear PRI 13-JAN-1995  
LOCUS Human statherin mRNA, complete cds.  
DEFINITION M18371  
ACCESSION M18371.1 GI:338610  
VERSION statherin.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 542)  
AUTHORS Dickinson,D.P., Ridall,A.L. and Levine,M.J.  
TITLE Human submandibular gland statherin and basic histidine-rich peptide are encoded by highly abundant mRNA's derived from a common ancestral sequence  
JOURNAL Biochem. Biophys. Res. Commun. 149 (2), 784-790 (1987)  
MEDLINE 88106506  
PUBMED 3426601  
COMMENT Original source text: Human female submandibular gland, cDNA to mRNA, clone pBRSMGF98.2.  
Draft entry and computer readable copy of sequence [1] kindly provided by D.P.Dickinson 21-MAR-1988.

## FEATURES

source  
1..542  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="4q11-q13"  
1..542  
/gene="STATH"  
57..245  
/gene="STATH"  
/note="statherin precursor"  
/codon\_start=1  
/protein\_id="AAA60600.1"  
/db\_xref="GI:338611"  
/db\_xref="GDB:G00-120-391"  
/translation="MKFLVFAFILALMVSMIGADSEKFLRIGRFGYGPYPVP"

## gene

## CDS

sig\_peptide  
57..113  
/gene="STATH"  
/note="statherin signal peptide"  
114..242  
/gene="STATH"  
/product="statherin"  
ORIGIN 262 bp upstream of PstI site; chromosome 4q11-q13.  
Query Match 11.2%; Score 67.6; DB 9; Length 542;  
Best Local Similarity 61.2%; Pred. No. 0.0044;  
Matches 147; Conservative 0; Mismatches 84; Indels 9; Gaps 2;  
QY 140 GGACTCCACCAATATGAAGATCTTTATCTTTGCTTTCATTTATGGCTCTCATCCCA 199  
DB 43 GAACCCAGCCAACTATGAAGTTCTTGTCTTGTCTTCATCTTGGCTCTCATGGTTTCCA 102  
QY 200 TCATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGAAACGGAAGAAACATCATAGAG 259  
DB 103 TGATTGGAGCTGATTCATCTGAAGAGAAATTTTCCGTAGAAATTTGGAAGATTCGGTTATG 162  
QY 260 GATA-----TTTTCACAAATACCAAGCCATATCAACGATATCCACTAAATATCCTCTG 313  
DB 163 GGTATGGCCCTTATCAGCAGTTCAGAAACACCACTATATCCACACCAACCAACAC 222  
QY 314 CGT---ATCCATTTCTTAAATGCTGCTTAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 370  
DB 223 AATACCAACAATATACCTTTTAATATCATCATCACTAGTAAGTAAGTAAGTAAGTAAG 282  
RESULT 14  
HUMSTATHA HUMSTATHA 552 bp mRNA linear PRI 13-JAN-1995  
LOCUS Human statherin mRNA, complete cds.  
DEFINITION M18078  
ACCESSION M18078.1 GI:338507  
VERSION regulatory protein; statherin.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 552)  
AUTHORS Sabatini,L.M., Carlock,L.R., Johnson,G.W. and Azen,E.A.  
TITLE cDNA cloning and chromosomal localization (4q11-13) of a gene for statherin, a regulator of calcium in saliva  
JOURNAL Am. J. Hum. Genet. 41 (6), 1048-1060 (1987)  
MEDLINE 88074310  
PUBMED 3502720  
COMMENT Original source text: Human parotid gland, cDNA to mRNA, clone H772B.  
Draft entry and computer-readable sequence [1] kindly submitted by L.Sabatini 19-JAN-1988.  
FEATURES  
source  
1..552  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="4q11-q13"  
1..552  
/gene="STATH"  
73..261  
/gene="STATH"  
/note="statherin protein"  
/codon\_start=1  
/protein\_id="AAA60594.1"  
/db\_xref="GI:338508"  
/db\_xref="GDB:G00-120-391"  
/translation="MKFLVFAFILALMVSMIGADSEKFLRIGRFGYGPYPVP"  
EOPLYPOPYQPOYYTF"  
73..129  
/gene="STATH"  
/note="statherin signal peptide"  
ORIGIN 278 bp upstream of PstI site; chromosome 4q11-q13.  
sig\_peptide



Query Match 11.2%; Score 67.6; DB 9; Length 552;  
Best Local Similarity 61.2%; Pred. No. 0.0044;  
Matches 147; Conservative 0; Mismatches 84; Indels 9; Gaps 2;

QY 140 GGACTCCACCAATATGAGATCTTTATCTTTTGTCTTCAATTATGGCTCTCATCTAGCCA 199  
Db 59 GAACCCAGCCAACTATGAAGTCTCTTTTGGCTTCATCTTGGCTCTCATGGTTTCCA 118  
QY 200 TGATTAGAGCTGATTCATCTGAGAGAAAGCTCACAGGAACGGAAACACATCATAGAG 259  
Db 119 TGATTGGAGCTGATTCATCTGAGAGAAATTTTGGCTAGAAATTGGAAGATTCGGTTATG 178  
QY 260 GATA-----TTTCAACAATACCAGCATATCAACGATATCCACTAAATATCTCTCTG 313  
Db 179 GGTATGGCCCTTATCAGCCAGTTCAGAACACACACTATACCCACACCATACCAACCAC 238  
QY 314 CGT---ATCCATTTCTTAAATGCTCTTAGTAACTACAGGACATGATTAGAGATTT 370  
Db 239 AATACCAACAATATACCTTTTAAATATCATCATGAACCTGACGAGCATGATTATTGAGGCTT 298

RESULT 15  
AX747668  
LOCUS AX747668 1584 bp mRNA linear PAT 20-JUN-2003  
DEFINITION Sequence 1193 from Patent EP1308459.  
ACCESSION AX747668  
VERSION AX747668.1 GI:32132056  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,  
Yamamoto, J. I., Isozaki, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,  
Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and  
Masuho, Y.  
TITLE Full-length cDNA sequences  
JOURNAL Patent: EP 1308459-A 1193 07-MAY-2003;  
Helix Research Institute (JP); Research Association for  
Biotechnology (JP)  
FEATURES  
source 1..1584  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 11.2%; Score 67.6; DB 6; Length 1584;  
Best Local Similarity 61.2%; Pred. No. 0.0037;  
Matches 147; Conservative 0; Mismatches 84; Indels 9; Gaps 2;

QY 140 GGACTCCACCAATATGAGATCTTTATCTTTTGTCTTCAATTATGGCTCTCATCTAGCCA 199  
Db 1083 GAACCCAGCCAACTATGAAGTCTCTTTTGGCTTCATCTTGGCTCTCATGGTTTCCA 1142  
QY 200 TGATTAGAGCTGATTCATCTGAGAGAAAGCTCACAGGAACGGAAACACATCATAGAG 259  
Db 1143 TGATTGGAGCTGATTCATCTGAGAGAAATTTTGGCTAGAAATTGGAAGATTCGGTTATG 1202  
QY 260 GATA-----TTTCAACAATACCAGCATATCAACGATATCCACTAAATATCTCTCTG 313  
Db 1203 GGTATGGCCCTTATCAGCCAGTTCAGAACACACACTATACCCACACCATACCAACCAC 1262  
QY 314 CGT---ATCCATTTCTTAAATGCTCTTAGTAACTACAGGACATGATTAGAGATTT 370  
Db 1263 AATACCAACAATATACCTTTTAAATATCATCATGAACCTGACGAGCATGATTATTGAGGCTT 1322

Search completed: August 15, 2004, 22:31:24  
Job time : 2774 secs

---

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 14, 2004, 05:08:11 ; Search time 352 Seconds

(without alignments)  
7289.522 Million cell updates/sec

Title: US-10-079-754A-4

Perfect score: 604

Sequence: 1 gaagatttcctgctctata.....tatcaagcagcaaaaaaaaaa 604

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_29Jan04:\*

- 1: Geneseq1990s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001as:\*
- 5: Geneseq2001bs:\*
- 6: Geneseq2002s:\*
- 7: Geneseq2003as:\*
- 8: Geneseq2003bs:\*
- 9: Geneseq2003cs:\*
- 10: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	604	100.0	604	7 ABX11391	Abx11391 Bovine hi
2	465.4	77.1	505	7 ABX11388	Abx11388 Bovine hi
3	448.2	74.2	585	7 ABX11389	Abx11389 Bovine hi
4	423.4	70.1	525	7 ABX11393	Abx11393 Bovine hi
5	238.4	39.5	267	7 ABX11394	Abx11394 Bovine ma
6	115	19.0	869	7 ABX11390	Abx11390 Bovine hi
7	87	14.4	96	7 ABX11392	Abx11392 Bovine hi
8	75.4	12.5	357	2 AAV89526	AAV89526 EST clone
9	70.4	11.7	438	7 ACC51062	Acc51062 Human Chi
10	70.4	11.7	438	7 ABZ36590	Abz36590 Human GEN
11	70.4	11.7	437	7 ABZ36425	Abz36425 Human GEN
12	68.2	11.3	857	4 AAH98658	AAH98658 Human EST
13	67.6	11.2	552	3 AAA48964	AAA48964 Human sta
14	67.6	11.2	678	9 ADE40261	Ade40261 Human NOV
15	67.6	11.2	1584	9 ADB63039	ADB63039 Human CDN
16	67.2	11.1	203	2 AAV89600	AAV89600 EST clone
17	62.6	10.4	491	3 AAA48965	AAA48965 Human bas
18	62	10.3	406	9 ADE09636	Ade09636 Novel DNA
19	60.8	10.1	417	9 ADE07409	Ade07409 Novel cod
20	60.8	10.1	516	2 AAV89525	AAV89525 EST clone
21	60.8	10.1	548	8 AAL60905	AAL60905 Human sal
22	59.6	9.9	732	8 AAL61238	AAL61238 Human sec
23	58	9.6	457	6 ABQ99552	ABQ99552 Human cod

#### ALIGNMENTS

RESULT 1

ABX11391

ID ABX11391 standard; cDNA; 604 BP.

XX AC ABX11391;

XX DT 01-MAY-2003 (first entry)

XX DE Bovine histatin like polypeptide, #4, cDNA.

XX KW Cow; gene; ss; histatin; bovine; salivary secretion; oral cavity;

XX KW non-immune defence system; oral candidosis; gene therapy;

XX KW microbial infection; fungal infection; dental caries; plaque; tartar;

XX KW cystic fibrosis; systemic infection; Candida infection; mastitis;

XX KW fungicide; antibacterial.

XX OS Bos taurus.

XX FH Key

XX CDS

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

Claim 1; Page 12; 17pp; English.

PS The invention discloses isolated histatin polynucleotides and  
XX polypeptides, which comprise sequences expressed in bovine. Histatins are  
CC histidine rich and specific to the salivary secretions. They are believed  
CC to function as part of the non-immune defence system, particularly in the  
CC oral cavity and have promise as therapeutic agents in humans with oral  
CC candidosis. Also disclosed are compositions comprising the histatin  
CC polypeptide or polynucleotide, and at least one component consisting of  
CC physiological or pharmaceutical carriers or immunostimulants. The  
CC histatin polynucleotide, polypeptide, or the cosmetic composition  
CC comprising the polypeptide, is useful for treating (e.g. gene therapy) a  
CC disorder in a mammal, particularly microbial or fungal infections. The  
CC histatin polynucleotide is also useful in genome mapping, physical  
CC mapping or in the positional cloning of genes. Specifically, the  
CC polypeptide or polynucleotide is useful for treating fungal or bacterial  
CC infections of the oral cavity (e.g. dental caries, plaque or tartar),  
CC vagina, urethra, ear, skin, respiratory tract (e.g. cystic fibrosis  
CC caused by *Pseudomonas aeruginosa*), mucosa or eye, as well as for treating  
CC systemic infections such as systemic *Candida* infection. The polypeptide  
CC or polynucleotide is also useful for veterinary applications e.g. for  
CC treating mastitis. The sequence presented is a cDNA encoding a bovine  
CC histatin like polypeptide

SQ Sequence 604 BP; 222 A; 112 C; 79 G; 191 T; 0 U; 0 Other;

Query Match 100.0%; Score 604; DB 7; Length 604;

Best Local Similarity 100.0%; Pred. No. 5.2e-120;  
Matches 604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GAATATTTTCAGTTCTATATAGATCTCATACCTGATGTAATTAACAAACAAATGAA	60
Db	1	GAATATTTTCAGTTCTATATAGATCTCATACCTGATGTAATTAACAAACAAATGAA	60
Qy	61	GGATTTCAAGTATTTAAACACAGCAGTTTCTAGCAAGAAACATCTCTGAAGCATCAG	120
Db	61	GGATTTCAAGTATTTAAACACAGCAGTTTCTAGCAAGAAACATCTCTGAAGCATCAG	120
Qy	121	AATTTTCATCTTCATGACTCGACTCCACCAAAATGAAGATCTTTATCTTTGCTTCATT	180
Db	121	AATTTTCATCTTCATGACTCGACTCCACCAAAATGAAGATCTTTATCTTTGCTTCATT	180
Qy	181	ATGCTCTCATCTCCAGCATGATTTAGAGCTGATTCATCTGAAGAGAAACGTCACAGGAA	240
Db	181	ATGCTCTCATCTCCAGCATGATTTAGAGCTGATTCATCTGAAGAGAAACGTCACAGGAA	240
Qy	241	CGGAAAAACATCATAGAGATATTTTCACAAATCCAGCCATCAACGATATCCACTA	300
Db	241	CGGAAAAACATCATAGAGATATTTTCACAAATCCAGCCATCAACGATATCCACTA	300
Qy	301	AATTTATCTCTCGGTATCCATTTCTTTAAATGCTGTTAGTAATACACAGGACATGATT	360
Db	301	AATTTATCTCTCGGTATCCATTTCTTTAAATGCTGTTAGTAATACACAGGACATGATT	360
Qy	361	AGAGAGATTTTCAATGATTTTCTACTCTTCTGTTGTTGAAACCATCTTTCA	420
Db	361	AGAGAGATTTTCAATGATTTTCTACTCTTCTGTTGTTGAAACCATCTTTCA	420
Qy	421	AATCAATTAACAAAGAAAAAATAATCAGTCAAGTAGTTGTCACAAACACATCTTGAATC	480
Db	421	AATCAATTAACAAAGAAAAAATAATCAGTCAAGTAGTTGTCACAAACACATCTTGAATC	480
Qy	481	AAATATCAATATTTTAAAAACATAATATGATGTTCTGACATATGTAATGTTTCTTAC	540
Db	481	AAATATCAATATTTTAAAAACATAATATGATGTTCTGACATATGTAATGTTTCTTAC	540
Qy	541	TTTCTTTTCTCTGTCATCTACCATGCTGTTAATTAATGATCTATCAAGCATATAAAA	600
Db	541	TTTCTTTTCTCTGTCATCTACCATGCTGTTAATTAATGATCTATCAAGCATATAAAA	600
Qy	601	AAAA 604	
Db	601	AAAA 604	

RESULT 2  
ABX11388  
ID ABX11388 standard; cDNA; 505 BP.  
XX  
AC ABX11388;  
XX  
DT 01-MAY-2003 (first entry)  
DE Bovine histatin like polypeptide, #1, cDNA.  
XX  
KW Cow; gene; ss; histatin; bovine; salivary secretion; oral cavity;  
KW non-immune defence system; oral candidosis; gene therapy;  
KW microbial infection; fungal infection; dental caries; plaque; tartar;  
KW cystic fibrosis; systemic infection; *Candida* infection; mastitis;  
KW fungicide; antibacterial.  
XX  
OS *Bos taurus*.  
XX  
FH Key Location/Qualifiers  
FT 44..220  
FT /\*tag= a  
FT /product= "Histatin like polypeptide"  
XX  
XX US2002164625-A1.  
XX  
XX 07-NOV-2002.  
XX  
XX 19-FEB-2002; 2002US-00079754.  
XX  
XX 23-AUG-1999; 99US-0150330P.  
PR 29-OCT-1999; 99US-0162701P.  
XX  
PR 22-AUG-2000; 2000US-00644190.  
XX  
PR 27-OCT-2000; 2000US-00699146.  
XX  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
XX Glenn M, Grigor MR, Molenaar AJ, Davis SR;  
XX  
XX WPI: 2003-275306/27.  
DR P-PSDB; ABG75791.  
XX  
XX New histatin polynucleotides and polypeptides expressed in bovine  
PT tissues, useful for treating bacterial or fungal infections in mammals,  
PT e.g. infections of the oral cavity, vagina, urethra, ear or skin, or  
PT systemic infections.  
XX  
XX Claim 1; Page 11; 17pp; English.  
XX  
XX The invention discloses isolated histatin polynucleotides and  
CC polypeptides, which comprise sequences expressed in bovine. Histatins are  
CC histidine rich and specific to the salivary secretions. They are believed  
CC to function as part of the non-immune defence system, particularly in the  
CC oral cavity and have promise as therapeutic agents in humans with oral  
CC candidosis. Also disclosed are compositions comprising the histatin  
CC polypeptide or polynucleotide, and at least one component consisting of  
CC physiological or pharmaceutical carriers or immunostimulants. The  
CC histatin polynucleotide, polypeptide, or the cosmetic composition  
CC comprising the polypeptide, is useful for treating (e.g. gene therapy) a  
CC disorder in a mammal, particularly microbial or fungal infections. The  
CC histatin polynucleotide is also useful in genome mapping, physical  
CC mapping or in the positional cloning of genes. Specifically, the  
CC polypeptide or polynucleotide is useful for treating fungal or bacterial  
CC infections of the oral cavity (e.g. dental caries, plaque or tartar),  
CC vagina, urethra, ear, skin, respiratory tract (e.g. cystic fibrosis  
CC caused by *Pseudomonas aeruginosa*), mucosa or eye, as well as for treating  
CC systemic infections such as systemic *Candida* infection. The polypeptide  
CC or polynucleotide is also useful for veterinary applications e.g. for  
CC treating mastitis. The sequence presented is a cDNA encoding a bovine  
CC histatin like polypeptide  
XX  
XX Sequence 505 BP; 188 A; 96 C; 67 G; 154 T; 0 U; 0 Other;



RESULT 4  
ABX11393  
ID ABX11393 standard; cDNA; 525 BP.  
AC AC  
XX AC ABX11393;  
XX  
XX 01-MAY-2003 (first entry)  
XX Bovine histatin like polypeptide, #6, cDNA.  
XX  
XX Cow; gene; ss; histatin; bovine; salivary secretion; oral cavity;  
KW non-immune defence system; oral candidosis; gene therapy;  
KW microbial infection; fungal infection; dental caries; plaque; tartar;  
KW cystic fibrosis; systemic infection; Candida infection; mastitis;  
KW fungicide; antibacterial.  
XX  
XX Bos taurus.  
OS  
XX  
XX Key Location/Qualifiers  
FH 43..222  
FT /\*tag= a  
FT /product= "Histatin like polypeptide"  
XX  
XX US2002164625-A1.  
XX  
XX 07-NOV-2002.  
XX  
XX 19-FEB-2002; 2002US-00079754.  
XX PF  
XX 23-AUG-1999; 99US-0150330P.  
XX PR 29-OCT-1999; 99US-0162701P.  
XX PR 22-AUG-2000; 2000US-00644190.  
XX PR 27-OCT-2000; 2000US-00699146.  
XX  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
XX Glenn M, Grigor MR, Molenaar AJ, Davis SR;  
XX WPI; 2003-275306/27.  
XX P-PSDB; ABG75796.  
XX  
XX New histatin polynucleotides and polypeptides expressed in bovine  
XX tissues, useful for treating bacterial or fungal infections in mammals,  
XX e.g. infections of the oral cavity, vagina, urethra, ear or skin, or  
XX systemic infections.  
XX  
XX Claim 1; Page 13; 17pp; English.  
XX  
XX The invention discloses isolated histatin polynucleotides and  
XX polypeptides, which comprise sequences expressed in bovine. Histatins are  
XX histidine rich and specific to the salivary secretions. They are believed  
XX to function as part of the non-immune defence system, particularly in the  
XX oral cavity and have promise as therapeutic agents in humans with oral  
XX candidosis. Also disclosed are compositions comprising the histatin  
XX polypeptide or polynucleotide, and at least one component consisting of  
XX physiological or pharmaceutical carriers or immunostimulants. The  
XX histatin polynucleotide, polypeptide, or the cosmetic composition  
XX comprising the polypeptide, is useful for treating (e.g. gene therapy) a  
XX disorder in a mammal, particularly microbial or fungal infections. The  
XX histatin polynucleotide is also useful in genome mapping, physical  
XX mapping or in the positional cloning of genes. Specifically, the  
XX polypeptide or polynucleotide is useful for treating fungal or bacterial  
XX infections of the oral cavity (e.g. dental caries, plaque or tartar),  
XX vagina, urethra, ear, skin, respiratory tract (e.g. cystic fibrosis  
XX caused by *Pseudomonas aeruginosa*), mucosa or eye, as well as for treating  
XX systemic infections such as systemic *Candida* infection. The polypeptide  
XX or polynucleotide is also useful for veterinary applications e.g. for  
XX treating mastitis. The sequence presented is a cDNA encoding a bovine  
XX histatin like polypeptide  
XX  
XX Sequence 525 BP; 187 A; 102 C; 72 G; 164 T; 0 U; 0 Other;

Query Match 70.1%; Score 423.4; DB 7; Length 525;  
Best Local Similarity 93.4%; Pred. No. 2.4e-81;  
Matches 466; Conservative 0; Mismatches 1; Indels 32; Gaps 1;  
QY 138 CTGGACTCCACCAATATGAAGATCTTTATCTTTGCTTCTTCAATATGCTCTCATCTAGC 197  
DB 27 CAGGACTCCACCAATATGAAGATCTTTATCTTTGCTTCTTCAATATGCTCTCATCTAGC 86  
QY 198 CATGATTAGAGCTGATTCTATCTGAAGAGAAACGTCACAGGAAACGGAAAAACATCAT-- 255  
DB 87 CATGATTAGAGCTGATTCTATCTGAAGAGAAACGTCACAGGAAACGGAAAAACATCATGT 146  
QY 256 -----AGAGGATATTTTCAACAATACCAAGCCATAT 285  
DB 147 TGATAGGTTCTCCAGAAATCTTACTAATACAGAGGATATTTTCAACAATACCAAGCCATAT 206  
QY 286 CAAGGATATCCACTAAATATATCTCTCGCTATCCATTTCTTTAAAAATGCTGTAGTAA 345  
DB 207 CAAGGATATCCACTAAATATATCTCTCGCTATCCATTTCTTTAAAAATGCTGTAGTAA 266  
QY 346 CTACAGGACATGATTAGAGAGATTTTTCACATGATTTTTCCTACTCTTTCTGTGTGTT 405  
DB 267 CTACAGGACATGATTAGAGAGATTTTTCACATGATTTTTCCTACTCTTTCTGTGTGTT 326  
QY 406 GAAACCAATCTTTCAAAATGAATAAAACAAAGAAAAAATTCAGTCAAGTAGTTGCACAA 465  
DB 327 GAAACCAATCTTTCAAAATGAATAAAACAAAGAAAAAATTCAGTCAAGTAGTTGCACAA 386  
QY 466 CACATACCTTGGAAATCAAAATATCAATATTTTAAACATAAATGATAGTCTTGAAGTAT 525  
DB 387 CACATACCTTGGAAATCAAAATATCAATATTTTAAACATAAATGATAGTCTTGAAGTAT 446  
QY 526 GTAATGGTTTCTACTTTCTTCTGTCATTTACCATGTCATGCTTTAAATAATTTGATCT 585  
DB 447 GTAATGGTTTCTACTTTCTTCTGTCATTTACCATGTCATGCTTTAAATAATTTGATCT 506  
QY 586 ATCAAGCATATAAAAAA 604  
DB 507 ATCAAGCATATAAAAAA 525  
RESULT 5  
ABX11394  
ID ABX11394 standard; cDNA; 267 BP.  
XX AC ABX11394;  
XX  
XX 01-MAY-2003 (first entry)  
XX  
XX Bovine mammary tissue cDNA clone.  
XX  
XX Cow; ss; histatin; bovine; salivary secretion; oral cavity;  
KW non-immune defence system; oral candidosis; gene therapy;  
KW microbial infection; fungal infection; dental caries; plaque; tartar;  
KW cystic fibrosis; systemic infection; Candida infection; mastitis;  
KW fungicide; antibacterial; mammary-gland.  
XX  
XX Bos taurus.  
XX  
XX US2002164625-A1.  
XX  
XX 07-NOV-2002.  
XX  
XX 19-FEB-2002; 2002US-00079754.  
XX PR 23-AUG-1999; 99US-0150330P.  
XX PR 29-OCT-1999; 99US-0162701P.  
XX PR 22-AUG-2000; 2000US-00644190.  
XX PR 27-OCT-2000; 2000US-00699146.  
XX  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
XX Glenn M, Grigor MR, Molenaar AJ, Davis SR;  
XX

XX WPI; 2003-275306/27.  
XX  
XX New histatin polynucleotides and polypeptides expressed in bovine  
PT tissues, useful for treating bacterial or fungal infections in mammals,  
PT e.g. infections of the oral cavity, vagina, urethra, ear or skin, or  
PT systemic infections.  
XX  
XX Example 2; Page 15; 17pp; English.  
XX  
XX The invention discloses isolated histatin polynucleotides and  
CC polypeptides, which comprise sequences expressed in bovine. Histatins are  
CC histidine rich and specific to the salivary secretions. They are believed  
CC to function as part of the non-immune defence system, particularly in the  
CC oral cavity and have promise as therapeutic agents in humans with oral  
CC candidosis. Also disclosed are compositions comprising the histatin  
CC polypeptide or polynucleotide, and at least one component consisting of  
CC physiological or pharmaceutical carriers or immunostimulants. The  
CC histatin polynucleotide, polypeptide, or the cosmetic composition  
CC comprising the polypeptide, is useful for treating (e.g. gene therapy) a  
CC disorder in a mammal, particularly microbial or fungal infections. The  
CC histatin polynucleotide is also useful in genome mapping, physical  
CC mapping or in the positional cloning of genes. Specifically, the  
CC polypeptide or polynucleotide is useful for treating fungal or bacterial  
CC infections of the oral cavity (e.g. dental caries, plaque or tartar),  
CC vagina, urethra, ear, skin, respiratory tract (e.g. cystic fibrosis,  
CC caused by Pseudomonas aeruginosa), mucosa or eye, as well as for treating  
CC systemic infections such as systemic Candida infection. The polypeptide  
CC or polynucleotide is also useful for veterinary applications e.g. for  
CC treating mastitis. The sequence presented is the bovine mammary tissue  
CC cDNA clone which was used to illustrate mRNA expression levels of mammary  
CC -gland specific sequences  
XX  
XX Sequence 267 BP; 89 A; 59 C; 43 G; 76 T; 0 U; 0 Other;

Query Match 39.5%; Score 238.4; DB 7; Length 267;  
Best Local Similarity 99.6%; Pred. No. 8.9e-42;  
Matches 239; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 138 CTGAGCTCCACCAATATGAGATCTTATCTTGTCTTCATATGCTCTCATCTAGC 197  
Db 28 CAGGACTCCACCAATATGAGATCTTATCTTGTCTTCATATGCTCTCATCTAGC 87  
QY 198 CATGATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGGAACGGAACCAATCATAG 257  
Db 88 CATGATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGGAACGGAACCAATCATAG 147  
QY 258 AGGATATTTTCAACATACAGGCAATATCAACGATATCCACTAAATATCTCTCGGTA 317  
Db 148 AGGATATTTTCAACATACAGGCAATATCAACGATATCCACTAAATATCTCTCGGTA 207  
QY 318 TCCATTTTCCCTTAAATGCTGTTAGTAACTACAGGACATGATTAGAGAGATTTTTCACAA 377  
Db 208 TCCATTTTCCCTTAAATGCTGTTAGTAACTACAGGACATGATTAGAGAGATTTTTCACAA 267

RESULT 6  
ABX11390  
ID ABX11390 standard; cDNA; 869 BP.  
XX  
XX ABX11390;  
AC  
XX  
XX 01-MAY-2003 (first entry)  
DT  
XX  
XX Bovine histatin like polypeptide, #3, cDNA.  
XX  
XX Cow; gene; ss; histatin; bovine; salivary secretion; oral cavity;  
XX non-immune defence system; oral candidosis; gene therapy;  
XX microbial infection; fungal infection; dental caries; plaque; tartar;  
XX cystic fibrosis; systemic infection; Candida infection; mastitis;  
XX fungicide; antibacterial.  
XX  
XX Bos taurus.

XX Key Location/Qualifiers  
FH CDS 72..284  
FT /\*tag= a  
FT /product= "Histatin like polypeptide"  
XX  
XX US2002164625-A1.  
XX  
XX 07-NOV-2002.  
XX  
XX 19-FEB-2002; 2002US-00079754.  
XX  
XX 23-AUG-1999; 99US-0150330P.  
XX 29-OCT-1999; 99US-0162701P.  
XX 22-AUG-2000; 2000US-00644190.  
XX 27-OCT-2000; 2000US-00699146.  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
XX Glenn M, Grigor MR, Molenaar AJ, Davis SR;  
PI WPI; 2003-275306/27.  
XX P-PSDB; ABG75793.  
XX  
XX New histatin polynucleotides and polypeptides expressed in bovine  
PT tissues, useful for treating bacterial or fungal infections in mammals,  
PT e.g. infections of the oral cavity, vagina, urethra, ear or skin, or  
PT systemic infections.  
XX  
XX Claim 1; Page 12; 17pp; English.  
XX  
XX The invention discloses isolated histatin polynucleotides and  
CC polypeptides, which comprise sequences expressed in bovine. Histatins are  
CC histidine rich and specific to the salivary secretions. They are believed  
CC to function as part of the non-immune defence system, particularly in the  
CC oral cavity and have promise as therapeutic agents in humans with oral  
CC candidosis. Also disclosed are compositions comprising the histatin  
CC polypeptide or polynucleotide, and at least one component consisting of  
CC physiological or pharmaceutical carriers or immunostimulants. The  
CC histatin polynucleotide, polypeptide, or the cosmetic composition  
CC comprising the polypeptide, is useful for treating (e.g. gene therapy) a  
CC disorder in a mammal, particularly microbial or fungal infections. The  
CC histatin polynucleotide is also useful in genome mapping, physical  
CC mapping or in the positional cloning of genes. Specifically, the  
CC polypeptide or polynucleotide is useful for treating fungal or bacterial  
CC infections of the oral cavity (e.g. dental caries, plaque or tartar),  
CC vagina, urethra, ear, skin, respiratory tract (e.g. cystic fibrosis,  
CC caused by Pseudomonas aeruginosa), mucosa or eye, as well as for treating  
CC systemic infections such as systemic Candida infection. The polypeptide  
CC or polynucleotide is also useful for veterinary applications e.g. for  
CC treating mastitis. The sequence presented is a cDNA encoding a bovine  
CC histatin like polypeptide  
XX  
XX Sequence 869 BP; 282 A; 157 C; 147 G; 283 T; 0 U; 0 Other;  
SQ  
Query Match 19.0%; Score 115; DB 7; Length 869;  
Best Local Similarity 92.4%; Pred. No. 2.9e-15;  
Matches 121; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 138 CTGAGCTCCACCAATATGAGATCTTATCTTGTCTTCATATGCTCTCATCTAGC 197  
Db 56 CAGGACTCCACCAATATGAGATCTTATCTTGTCTTCATATGCTCTCATCTAGC 115  
QY 198 CATGATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGGAACGGAACCAATCATAG 257  
Db 116 CATGATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGGAACGGAACCAATCATAG 175  
QY 258 AGGATATTTTTC 268  
Db 176 ATGATTCTCTC 186  
XX  
XX RESULT 7

Query Match 14.4%; Score 87; DB 7; Length 96;  
 Best Local Similarity 94.7%; Pred. No. 2e-09;  
 Matches 90; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 265 TTTCACAAATACAGCCATATCAAGCATATCCACTAAATATATCTCTCGGTATCCATTT 324  
 DB 1 TTTCACAAATACAGCCATATCAAGCATATCCACTAAATATATCTCTCGGTATCCATTT 60

QY 325 CCTTAAATGCTGCTTAGTAACACAGGACATGAT 359  
 DB 61 TCTTAAATGCTGCTTAGTAACACAGGACATGAT 95

RESULT 8  
 ID AAV89526 standard; cDNA: 357 BP.  
 AC AAV89526;  
 DT 15-FEB-1999 (first entry)  
 DE EST clone CP294.  
 KW Human; secreted protein; expressed sequence tag; EST; haematopoiesis;  
 KW tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;  
 KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;  
 KW gene therapy; ss.  
 OS Homo sapiens.  
 PN WC9845436-A2.  
 PD 15-OCT-1998.  
 PF 10-APR-1998; 98WO-US006955.  
 PR 10-APR-1997; 97US-00838821.  
 PA (GENE) GENETICS INST INC.  
 PI Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;  
 PI Spaulding V, Agostino MJ;  
 PI WPI; 1999-070077/06.  
 PT New polynucleotides encoding human secreted proteins - derived from e.g.  
 PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,  
 PT pituitary, retina and colon cDNA libraries.  
 PS Claim 1; Page 245; 618pp; English.

XX The present sequence represents a human expressed sequence tag (EST). The  
 CC polynucleotide, which is a secreted EST, and the encoded protein are  
 CC predicted to have useful biological activities which would make them  
 CC suitable for treating, preventing or ameliorating medical conditions in  
 CC humans and animals, although no supporting data is given. Suggested  
 CC activities include nutritional activity, immune stimulating or  
 CC suppressing activity, haematopoiesis regulating activity, tissue growth  
 CC activity, activin/inhibin activity, chemotactic/chemokinetic activity,  
 CC haemostatic and thrombolytic activity, receptor/ligand activity, anti-  
 CC inflammatory activity, cadherin/tumour, invasion suppressor activity,  
 CC tumour inhibition activity. The polynucleotide may also be useful for  
 CC gene therapy  
 CC Sequence 357 BP; 101 A; 74 C; 65 G; 117 T; 0 U; 0 Other;  
 SQ

Query Match 12.5%; Score 75.4; DB 2; Length 357;  
 Best Local Similarity 59.4%; Pred. No. 7.6e-07;  
 Matches 149; Conservative 0; Mismatches 96; Indels 6; Gaps 1;  
 QY 120 GAATTTTCATCTTTTCATGACTGCACCTCCACCAATATGAAGATCTTTATCTTTGCTTCAT 179  
 DB 48 GACTCTCTCTTGGAGTAAAGGACTCAGCAACTATGAAGTTTGTCTTGTCTTAGT 107

XX The invention discloses isolated histatin polynucleotides and  
 CC polypeptides, which comprise sequences expressed in bovine. Histatins are  
 CC histidine rich and specific to the salivary secretions. They are believed  
 CC to function as part of the non-immune defence system, particularly in the  
 CC oral cavity and have promise as therapeutic agents in humans with oral  
 CC candidosis. Also disclosed are compositions comprising the histatin  
 CC polypeptide or polynucleotide, and at least one component consisting of  
 CC physiological or pharmaceutical carriers or immunostimulants. The  
 CC histatin polynucleotide, polypeptide, or the cosmetic composition  
 CC comprising the polypeptide, is useful for treating (e.g. gene therapy) a  
 CC disorder in a mammal, particularly microbial or fungal infections. The  
 CC histatin polynucleotide is also useful in genome mapping, physical  
 CC mapping or in the positional cloning of genes. Specifically, the  
 CC polypeptide or polynucleotide is useful for treating fungal or bacterial  
 CC infections of the oral cavity (e.g. dental caries, plaque or tartar),  
 CC vagina, urethra, ear, skin, respiratory tract (e.g. cystic fibrosis),  
 CC caused by Pseudomonas aeruginosa, mucosa or eye, as well as for treating  
 CC systemic infections such as systemic Candida infection. The polypeptide  
 CC or polynucleotide is also useful for veterinary applications e.g. for  
 CC treating mastitis. The sequence presented is a cDNA encoding a bovine  
 CC histatin like polypeptide  
 CC Sequence 96 BP; 33 A; 25 C; 10 G; 28 T; 0 U; 0 Other;  
 SQ

XX New histatin polynucleotides and polypeptides expressed in bovine  
 PT tissues, useful for treating bacterial or fungal infections in mammals,  
 PT e.g. infections of the oral cavity, vagina, urethra, ear or skin, or  
 PT systemic infections.  
 XX Claim 1; Page 12; 17pp; English.

XX The invention discloses isolated histatin polynucleotides and  
 CC polypeptides, which comprise sequences expressed in bovine. Histatins are  
 CC histidine rich and specific to the salivary secretions. They are believed  
 CC to function as part of the non-immune defence system, particularly in the  
 CC oral cavity and have promise as therapeutic agents in humans with oral  
 CC candidosis. Also disclosed are compositions comprising the histatin  
 CC polypeptide or polynucleotide, and at least one component consisting of  
 CC physiological or pharmaceutical carriers or immunostimulants. The  
 CC histatin polynucleotide, polypeptide, or the cosmetic composition  
 CC comprising the polypeptide, is useful for treating (e.g. gene therapy) a  
 CC disorder in a mammal, particularly microbial or fungal infections. The  
 CC histatin polynucleotide is also useful in genome mapping, physical  
 CC mapping or in the positional cloning of genes. Specifically, the  
 CC polypeptide or polynucleotide is useful for treating fungal or bacterial  
 CC infections of the oral cavity (e.g. dental caries, plaque or tartar),  
 CC vagina, urethra, ear, skin, respiratory tract (e.g. cystic fibrosis),  
 CC caused by Pseudomonas aeruginosa, mucosa or eye, as well as for treating  
 CC systemic infections such as systemic Candida infection. The polypeptide  
 CC or polynucleotide is also useful for veterinary applications e.g. for  
 CC treating mastitis. The sequence presented is a cDNA encoding a bovine  
 CC histatin like polypeptide  
 CC Sequence 96 BP; 33 A; 25 C; 10 G; 28 T; 0 U; 0 Other;  
 SQ

XX New histatin polynucleotides and polypeptides expressed in bovine  
 PT tissues, useful for treating bacterial or fungal infections in mammals,  
 PT e.g. infections of the oral cavity, vagina, urethra, ear or skin, or  
 PT systemic infections.  
 XX Claim 1; Page 12; 17pp; English.

XX The invention discloses isolated histatin polynucleotides and  
 CC polypeptides, which comprise sequences expressed in bovine. Histatins are  
 CC histidine rich and specific to the salivary secretions. They are believed  
 CC to function as part of the non-immune defence system, particularly in the  
 CC oral cavity and have promise as therapeutic agents in humans with oral  
 CC candidosis. Also disclosed are compositions comprising the histatin  
 CC polypeptide or polynucleotide, and at least one component consisting of  
 CC physiological or pharmaceutical carriers or immunostimulants. The  
 CC histatin polynucleotide, polypeptide, or the cosmetic composition  
 CC comprising the polypeptide, is useful for treating (e.g. gene therapy) a  
 CC disorder in a mammal, particularly microbial or fungal infections. The  
 CC histatin polynucleotide is also useful in genome mapping, physical  
 CC mapping or in the positional cloning of genes. Specifically, the  
 CC polypeptide or polynucleotide is useful for treating fungal or bacterial  
 CC infections of the oral cavity (e.g. dental caries, plaque or tartar),  
 CC vagina, urethra, ear, skin, respiratory tract (e.g. cystic fibrosis),  
 CC caused by Pseudomonas aeruginosa, mucosa or eye, as well as for treating  
 CC systemic infections such as systemic Candida infection. The polypeptide  
 CC or polynucleotide is also useful for veterinary applications e.g. for  
 CC treating mastitis. The sequence presented is a cDNA encoding a bovine  
 CC histatin like polypeptide  
 CC Sequence 96 BP; 33 A; 25 C; 10 G; 28 T; 0 U; 0 Other;  
 SQ

XX New histatin polynucleotides and polypeptides expressed in bovine  
 PT tissues, useful for treating bacterial or fungal infections in mammals,  
 PT e.g. infections of the oral cavity, vagina, urethra, ear or skin, or  
 PT systemic infections.  
 XX Claim 1; Page 12; 17pp; English.

XX The invention discloses isolated histatin polynucleotides and  
 CC polypeptides, which comprise sequences expressed in bovine. Histatins are  
 CC histidine rich and specific to the salivary secretions. They are believed  
 CC to function as part of the non-immune defence system, particularly in the  
 CC oral cavity and have promise as therapeutic agents in humans with oral  
 CC candidosis. Also disclosed are compositions comprising the histatin  
 CC polypeptide or polynucleotide, and at least one component consisting of  
 CC physiological or pharmaceutical carriers or immunostimulants. The  
 CC histatin polynucleotide, polypeptide, or the cosmetic composition  
 CC comprising the polypeptide, is useful for treating (e.g. gene therapy) a  
 CC disorder in a mammal, particularly microbial or fungal infections. The  
 CC histatin polynucleotide is also useful in genome mapping, physical  
 CC mapping or in the positional cloning of genes. Specifically, the  
 CC polypeptide or polynucleotide is useful for treating fungal or bacterial  
 CC infections of the oral cavity (e.g. dental caries, plaque or tartar),  
 CC vagina, urethra, ear, skin, respiratory tract (e.g. cystic fibrosis),  
 CC caused by Pseudomonas aeruginosa, mucosa or eye, as well as for treating  
 CC systemic infections such as systemic Candida infection. The polypeptide  
 CC or polynucleotide is also useful for veterinary applications e.g. for  
 CC treating mastitis. The sequence presented is a cDNA encoding a bovine  
 CC histatin like polypeptide  
 CC Sequence 96 BP; 33 A; 25 C; 10 G; 28 T; 0 U; 0 Other;  
 SQ

XX New histatin polynucleotides and polypeptides expressed in bovine  
 PT tissues, useful for treating bacterial or fungal infections in mammals,  
 PT e.g. infections of the oral cavity, vagina, urethra, ear or skin, or  
 PT systemic infections.  
 XX Claim 1; Page 12; 17pp; English.

XX The invention discloses isolated histatin polynucleotides and  
 CC polypeptides, which comprise sequences expressed in bovine. Histatins are  
 CC histidine rich and specific to the salivary secretions. They are believed  
 CC to function as part of the non-immune defence system, particularly in the  
 CC oral cavity and have promise as therapeutic agents in humans with oral  
 CC candidosis. Also disclosed are compositions comprising the histatin  
 CC polypeptide or polynucleotide, and at least one component consisting of  
 CC physiological or pharmaceutical carriers or immunostimulants. The  
 CC histatin polynucleotide, polypeptide, or the cosmetic composition  
 CC comprising the polypeptide, is useful for treating (e.g. gene therapy) a  
 CC disorder in a mammal, particularly microbial or fungal infections. The  
 CC histatin polynucleotide is also useful in genome mapping, physical  
 CC mapping or in the positional cloning of genes. Specifically, the  
 CC polypeptide or polynucleotide is useful for treating fungal or bacterial  
 CC infections of the oral cavity (e.g. dental caries, plaque or tartar),  
 CC vagina, urethra, ear, skin, respiratory tract (e.g. cystic fibrosis),  
 CC caused by Pseudomonas aeruginosa, mucosa or eye, as well as for treating  
 CC systemic infections such as systemic Candida infection. The polypeptide  
 CC or polynucleotide is also useful for veterinary applications e.g. for  
 CC treating mastitis. The sequence presented is a cDNA encoding a bovine  
 CC histatin like polypeptide  
 CC Sequence 96 BP; 33 A; 25 C; 10 G; 28 T; 0 U; 0 Other;  
 SQ

XX New histatin polynucleotides and polypeptides expressed in bovine  
 PT tissues, useful for treating bacterial or fungal infections in mammals,  
 PT e.g. infections of the oral cavity, vagina, urethra, ear or skin, or  
 PT systemic infections.  
 XX Claim 1; Page 12; 17pp; English.



QY 180 TATGGCTCTCATCTAGCCAGTANTAGAGCTGATTCATCTGAAGAGAAACGTCACAGGAA 239  
Db 108 CTGGCTCTCATGATTTCCATGATTAGCGCTGATTCACATGAAGAGACATCATGGGTA 167  
QY 240 ACGGAAAAACATCATAGAGGATATTTCAACAATACCAGCCATATCAACGATATCCACT 299  
Db 168 TAGAGAAATTCAT-----GAAAGCATCATTCACATCGAGATTTCCATTTATGG 221  
QY 300 AAATTAATCCCTCGCGTATCCATTTCTTTAAATGCTGTAGTAATACTACAGACATGAT 359  
Db 222 GGAATGTGGATCAAAATATCTATATGACAATTCATATCTCTTAGTAATCATGGGCAATGAT 281  
QY 360 TAGAGATTT 370  
Db 282 TATAGAGTTT 292

RESULT 9  
ACC51062  
ID ACC51062 standard; cDNA; 438 BP.  
XX  
AC ACC51062;  
XX  
DT 13-JUN-2003 (first entry)  
XX  
DE Human Chimera coding sequence.  
XX  
KW Human; GENSET; therapeutic; therapy; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO200294864-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 06-AUG-2001; 2001WO-1B001715.  
XX  
PR 25-MAY-2001; 2001US-0293574P.  
PR 15-JUN-2001; 2001US-0298698P.  
PR 29-JUN-2001; 2001US-0302277P.  
PR 13-JUL-2001; 2001US-0305456P.  
XX  
PA (GEST ) GENSET.  
XX  
PI Bejanin S, Tanaka H;  
XX  
DR WPI; 2003-129412/12.  
DR P-PSDB; ABR48455.  
XX  
PT New GENSET polynucleotides and polypeptides, useful for preparing a  
PT composition for treating GENSET-related disorders and as reagents in  
PT assays to quantitatively determined levels of GENSET expression in  
PT biological samples.  
XX  
PS Claim 1; Page 404-405; 505pp; English.  
XX  
CC The present invention relates to novel human GENSET coding sequences  
CC (ACC51060-ACC51115) and proteins (ABR48453-ABR48508). The GENSET  
CC sequences are useful for preparing a composition for treating GENSET-  
CC related disorders. They can also be used as markers for tissues in which  
CC the corresponding protein is preferentially expressed, as molecular  
CC weight markers on Southern gels, as chromosome markers or tags to  
CC identify chromosomes, and as reagents in assays to quantitatively  
CC determined levels of GENSET expression in biological samples  
XX  
SQ Sequence 438 BP; 155 A; 80 C; 60 G; 143 T; 0 U; 0 Other;  
Query Match 11.7%; Score 70.4; DB 7; Length 438;  
Best Local Similarity 69.9%; Pred. No. 9.3e-06;  
Matches 95; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 120 GAATTTTCATCTTTTCATGCTGGACTCCACCAATATGAAGATCTTTATCTTTGCTTCAT 179

Db 50 GACTCTCTCTTGAGTAAAGGACTCAGCCAACTATGAAGTTTTTTGCTTTGCTTTAGT 109  
QY 180 TATGGCTCTCATCTAGCCAGTANTAGAGCTGATTCATCTGAAGAGAAACGTCACAGGAA 239  
Db 110 CTGGCTCTCATGATTTCCATGATTAGCGCTGATTCACATGAAGAGACATCATGGGTA 169  
QY 240 ACGGAAAAACATCAT 255  
Db 170 TAGAAGAAATTCAT 185

RESULT 10  
ABZ36590  
ID ABZ36590 standard; cDNA; 438 BP.  
XX  
AC ABZ36590;  
XX  
DT 21-FEB-2003 (first entry)  
XX  
DE Human GENSET coding sequence SEQ ID 358.  
XX  
KW Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;  
KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;  
KW inflammatory disease; immune disorder; neuromuscular; toxicity;  
KW central nervous system; cardiovascular; gastrointestinal; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO200283899-A1.  
XX  
PD 24-OCT-2002.  
XX  
PF 18-APR-2001; 2001WO-1B000914.  
XX  
PR 18-APR-2001; 2001WO-1B000914.  
XX  
PA (GEST ) GENSET.  
XX  
PI Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;  
XX  
DR WPI; 2003-075548/07.  
XX  
PT New GENSET polynucleotides and polypeptides, useful for treating heavy  
PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the  
PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the  
PT toxicity.  
XX  
PS Claim 12; Page 478; 735pp; English.  
XX  
CC The present invention relates to novel GENSET polynucleotides (ABZ36404-  
CC ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides  
CC and polypeptides are useful in screening and diagnostic assays for  
CC abnormal GENSET expression and/or biological activity. They are also  
CC useful for screening of compounds for treating or preventing GENSET-  
CC related disorders, such as heavy metal toxicity, cancer, inflammatory  
CC diseases, immune disorders, and the neuromuscular, central nervous system  
CC (CNS), cardiovascular or gastrointestinal effects of the toxicity. The  
CC polynucleotides are useful for constructing or expanding chromosome maps  
XX  
SQ Sequence 438 BP; 151 A; 80 C; 60 G; 145 T; 0 U; 2 Other;  
Query Match 11.7%; Score 70.4; DB 7; Length 438;  
Best Local Similarity 69.9%; Pred. No. 9.3e-06;  
Matches 95; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 120 GAATTTTCATCTTTTCATGCTGGACTCCACCAATATGAAGATCTTTATCTTTGCTTCAT 179  
Db 50 GACTCTCTCTTGAGTAAAGGACTCAGCCAACTATGAAGTTTTTTGCTTTGCTTTAGT 109  
QY 180 TATGGCTCTCATCTAGCCAGTANTAGAGCTGATTCATCTGAAGAGAAACGTCACAGGAA 239  
Db 110 CTGGCTCTCATGATTTCCATGATTAGCGCTGATTCACATGAAGAGACATCATGGGTA 169

QY 240 ACGGAAAAAACATCAT 255  
 DB 170 TAGAAGAAATTCAT 185  
  
 RESULT 11  
 ID ABZ36425 standard; cDNA; 438 BP.  
 AC ABZ36425;  
 XX 21-FEB-2003 (first entry)  
 DT Human GENSET coding sequence SEQ ID 22.  
 DE Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;  
 KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;  
 KW inflammatory disease; immune disorder; neuromuscular; toxicity;  
 KW central nervous system; cardiovascular; gastrointestinal; gene; ss.  
 XX Homo sapiens.  
 OS WO200283898-A1.  
 PN 24-OCT-2002.  
 PD 18-APR-2001; 2001WO-IB000914.  
 PF 18-APR-2001; 2001WO-IB000914.  
 PR (GEST ) GENSET.  
 PA Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;  
 PI WPI; 2003-075548/07.  
 XX New GENSET polynucleotides and polypeptides, useful for treating heavy  
 DR metal toxicity, cancer, inflammatory diseases, immune disorders, and the  
 XX neuromuscular, CNS, cardiovascular or gastrointestinal effects of the  
 PT toxicity.  
 PS Claim 12; Page 295; 735pp; English.  
 CC The present invention relates to novel GENSET polynucleotides (ABZ36404-  
 CC ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides  
 CC and polypeptides are useful in screening and diagnostic assays for  
 CC abnormal GENSET expression and/or biological activity. They are also  
 CC useful for screening of compounds for treating or preventing GENSET-  
 CC related disorders, such as heavy metal toxicity, cancer, inflammatory  
 CC diseases, immune disorders, and the neuromuscular, central nervous system  
 CC (CNS), cardiovascular or gastrointestinal effects of the toxicity. The  
 CC polynucleotides are useful for constructing or expanding chromosome maps  
 XX Sequence 438 BP; 153 A; 80 C; 60 G; 143 T; 0 U; 2 Other;  
 SQ  
 Query Match 11.7%; Score 70.4; DB 7; Length 438;  
 Best Local Similarity 69.9%; Pred. No. 9.3e-06;  
 Matches 95; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
 QY 120 GAATTCATCTTCATGACTGCTCCACCAATATGAAGATCTTTATCTTTGCTTCAT 179  
 DB 50 GACTCTCTCTTGAGTAAAGAGCTCAGCCAACTATGAAGTTTTCCTTTGCTTTAGT 109  
 QY 180 TATGGCTCTCATCTAGCCATGATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGAA 239  
 DB 110 CTGGCTCTCATGATTTCCATGATTAGGCTGATTCATCAATGAAAGAGACATCATGGTA 169  
 QY 240 ACGGAAAAAACATCAT 255  
 DB 170 TAGAAGAAATTCAT 185  
  
 RESULT 12  
 ID AAH98658 standard; cDNA; 857 BP.  
 AC AAH98658;  
 XX 12-OCT-2001 (first entry)  
 DT Human EST-derived coding sequence SEQ ID NO: 515.  
 DE Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
 KW diagnostics; forensic test; gene mapping; genetic disorder; biodiversity;  
 KW gene therapy; nutrition; ss.  
 XX Homo sapiens.  
 OS WO200154477-A2.  
 PN 02-AUG-2001.  
 PD 25-JAN-2001; 2001WO-US002687.  
 PF 25-JAN-2001; 2000US-00491404.  
 PR 17-JUL-2000; 2000US-00617746.  
 PR 03-AUG-2000; 2000US-00631451.  
 PR 15-SEP-2000; 2000US-00663870.  
 XX (HYSE-) HYSEQ INC.  
 PA Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;  
 XX WPI; 2001-476164/51.  
 DR P-PSDB; AAM23999.  
 XX Isolated polypeptide for treatment of diseases, diagnostics, raising  
 PT antibodies and research use.  
 PS Claim 1; Page 537; 1275pp; English.  
 CC The present invention provides the protein and coding sequences of novel  
 CC proteins from a variety of organisms, including human, dog, cat, horse,  
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
 CC from the organism of interest. They can be used in diagnostics,  
 CC forensics, gene mapping, identification of mutations, to assess  
 CC biodiversity and for nutritional purposes. The present sequence is a cDNA  
 CC of the invention  
 XX Sequence 857 BP; 231 A; 194 C; 169 G; 263 T; 0 U; 0 Other;  
 SQ  
 Query Match 11.3%; Score 68.2; DB 4; Length 857;  
 Best Local Similarity 65.4%; Pred. No. 3.1e-05;  
 Matches 100; Conservative 0; Mismatches 53; Indels 0; Gaps 0;  
 QY 120 GAATTCATCTTCATGACTGCTCCACCAATATGAAGATCTTTATCTTTGCTTCAT 179  
 DB 361 GATTCTCTCTTGAGTAAAGAGCTCAGCCAACTATGAAGTTTTCCTTTGCTTTAAT 420  
 QY 180 TATGGCTCTCATCTAGCCATGATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGAA 239  
 DB 421 CTGGCTCTCATGCTTTCCATGACTGGAGCTGATTCATCAATGAAAGAGACATCATGGTA 480  
 QY 240 ACGGAAAAAACATCATAGAGGATATTTCAACA 272  
 DB 481 TAAAGAAAAATTCATGAAAGACATCATTCACA 513  
  
 RESULT 13  
 ID AAA48964 standard; DNA; 552 BP.  
 ID AAA48964  
 XX

AC AAA48964;  
XX 06-OCT-2000 (first entry)  
XX Human statherin DNA.  
XX Lysine-rich statherin protein; LRSP; acidic phosphoprotein; human;  
KW precipitation inhibitor; autoimmune; inflammatory disorder; AIDS; asthma;  
KW allergy; diabetes mellitus; fungal; bacterial infection; cancer;  
KW leukemia; adenocarcinoma; melanoma; ds.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
XX CDS 73..281  
XX /\*tag= a  
XX /product= "Statherin"  
XX WO200024779-A1.  
XX 04-MAY-2000.  
XX 22-OCT-1999; 99WO-US024046.  
XX 23-OCT-1998; 98US-0155209P.  
XX (INCY-) INCYTE PHARM INC.  
XX Tang YT, Corley NC, Guegler KJ, Patterson C;  
XX WPI; 2000-350699/30.  
XX P-PSDB; AAY94527.  
XX Purified polypeptide used for treating or preventing a disorder  
XX characterized by expression or activity of lysine-rich statherin  
XX proteins.  
XX Disclosure; Page 71; 75pp; English.  
XX The present invention relates to human lysine-rich statherin protein  
XX (LRSP) (AAY94526). The cDNA sequence encoding this protein was identified  
XX through analysis of a cDNA library of breast tumour tissue (BRSTNCT14).  
XX The LRSP sequence was found to have homology with the DNA of human  
XX statherin (the present sequence) and human basic histidine-rich protein  
XX (AAY94528). Human statherin is a phosphoprotein that acts as an inhibitor  
XX of precipitation of calcium phosphate salts in the oral cavity. The LRSP  
XX polypeptide and its antagonists may be useful for treating or preventing  
XX disorders associated with the activity of LRSP. Such disorders include  
XX autoimmune/inflammatory disorders (for example AIDS, allergies, asthma,  
XX diabetes mellitus), bacterial and fungal infection and cancers (such as  
XX leukemia, adenocarcinoma, melanoma). Antibodies to LRSP may be useful for  
XX diagnosis of the above disorders  
XX Sequence 552 BP; 172 A; 122 C; 78 G; 180 T; 0 U; 0 Other;  
Query Match 11.2%; Score 67.6; DB 3; Length 552;  
Best Local Similarity 61.2%; Pred. No. 3.8e-05;  
Matches 147; Conservative 0; Mismatches 84; Indels 9; Gaps 2;  
Qy 140 GGATCTCACCAAAATATGAAGATCTTTATCTTTGCTTCATATGGCTCTCATCTAGCCA 199  
Db 59 GAACCCAGCCAACTATGAGTCTCTGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 118  
Qy 200 TGATTAGCTGATCTATCTGAAGAGAAACGTCACAGAAACCGGAAACAAATCATATAG 259  
Db 119 TGATTGGAGCTGATCTATCTGAAGAGAAATTTTTCGTAGATTTGGAAGATTCGGTTATG 178  
Qy 260 GATA-----TTTTCAACAATACAGCCATATCAAGATATCCACTAAATTCCTCCCTG 313  
Db 179 GGTATGCCCTTATCAGCCAGTTCAGAAACACCACTATACCCACACCAATACCAACAC 238  
Qy 314 CGT---ATCCATTCCTTAAATGCTGCTTAGTAACCTACAGGACATGATTAGAGATTT 370

Db 239 AATACCAACAATATACCTTTTAAATATCATCATCACTCACTCACTCACTCACTCACT 298  
RESULT 14  
ADE40261  
ID ADE40261 standard; cDNA; 678 BP.  
XX ADE40261;  
AC ADE40261;  
XX 29-JAN-2004 (first entry)  
DT Human NOV38a cDNA - SEQ ID 167.  
XX DE  
XX NOVX; cardiant; antiarteriosclerotic; hypotensive; cytostatic; anorectic;  
KW antidiabetic; immunosuppressive; anti-HIV; neuroprotective; nootropic;  
KW antiparkinsonian; antiasthmatic; gynaecological; cardiomyopathy;  
KW atherosclerosis; hypertension; cancer; obesity; diabetes; AIDS;  
KW multiple sclerosis; graft-versus-host disease; Alzheimer's; Parkinson's;  
KW asthma; fertility disorder; vaccine; gene therapy; chromosome mapping;  
KW tissue typing; human; NOV; ss; gene.  
XX Homo sapiens.  
XX WO2003064589-A2.  
XX 07-AUG-2003.  
XX 02-AUG-2002; 2002WO-US024483.  
XX 02-AUG-2001; 2001US-0309501P.  
XX 03-AUG-2001; 2001US-0310291P.  
XX 07-AUG-2001; 2001US-0310544P.  
XX 08-AUG-2001; 2001US-0310951P.  
XX 09-AUG-2001; 2001US-0311292P.  
XX 13-AUG-2001; 2001US-0311979P.  
XX 16-AUG-2001; 2001US-0312892P.  
XX 17-AUG-2001; 2001US-0313201P.  
XX 17-AUG-2001; 2001US-0313415P.  
XX 20-AUG-2001; 2001US-0313643P.  
XX 20-AUG-2001; 2001US-0313702P.  
XX 21-AUG-2001; 2001US-0314031P.  
XX 23-AUG-2001; 2001US-0314466P.  
XX 28-AUG-2001; 2001US-0315403P.  
XX 29-AUG-2001; 2001US-0315853P.  
XX 17-SEP-2001; 2001US-0322716P.  
XX 21-SEP-2001; 2001US-0323994P.  
XX 14-DEC-2001; 2001US-0340233P.  
XX 05-FEB-2002; 2002US-0354591P.  
XX 19-MAR-2002; 2002US-0365478P.  
XX 19-APR-2002; 2002US-0373814P.  
XX 19-APR-2002; 2002US-0373825P.  
XX 19-APR-2002; 2002US-0373989P.  
XX 23-APR-2002; 2002US-0374632P.  
XX 07-JUN-2002; 2002US-0386971P.  
XX 01-AUG-2002; 2002US-00210172.  
XX (CURA-) CURAGEN CORP.  
XX Kekuda R, Miller CE, Patturajan M, Pena CEA, Rieger DK;  
PI Shinkets RA, Zerhusen BD, Li L, Ji W, Padigaru M, Casman SU;  
PI Voss EZ, Boldog FL, Gorman L, Leite MW, Vernet CAM, Anderson DW;  
PI Guo X, Zhong M, Gerlach VL, Hjalte T, Rastelli L, Spytek KA;  
PI Edinger SR, Ellerman K, Malyankar UM, Macdougall JR, Stone DJ;  
PI Alsobrook JP, Lepley DW, Burgess CE, Majumder K, Wolenc AR;  
PI Smithson G;  
XX WPI; 2003-663472/62.  
XX P-PSDB; ADE40262.  
XX New NOVX polypeptides and nucleic acids, useful for preventing or  
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,  
PT atherosclerosis or diabetes, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.

XX Claim 20; SEQ ID NO 167; 560pp; English.  
 XX The invention relates to a novel NOVX polypeptide. The polypeptide of the  
 CC invention demonstrates cardiant, antiarteriosclerotic, hypotensive,  
 CC cyostatic, anorectic, antidiabetic, immunosuppressive, anti-HIV,  
 CC neuroprotective, nootropic, antiparkinsonian, antiasthmatic and  
 CC gynaeological activities and may be useful in diagnosing, treating or  
 CC preventing NOVX-associated disorders including cardiomyopathy,  
 CC atherosclerosis, hypertension, cancer, obesity, diabetes, Parkinson's  
 CC sclerosis, graft-versus-host disease, Alzheimer's disease, Parkinson's  
 CC disease, asthma or fertility disorders. Furthermore, the polypeptides may  
 CC be utilised as vaccines whilst the nucleic acids may be used as  
 CC hybridisation probes, in gene therapy, chromosome mapping, tissue typing,  
 CC preventive medicine and pharmacogenomics. The current sequence is that of  
 CC the human NOV cDNA of the invention.  
 XX  
 SQ Sequence 678 BP; 190 A; 151 C; 137 G; 200 T; 0 U; 0 Other;

Query Match 11.2%; Score 67.6; DB 9; Length 678;  
 Best Local Similarity 61.2%; Pred. No. 4e-05;  
 Matches 147; Conservative 0; Mismatches 84; Indels 9; Gaps 2;  
 140 GGACTCCACCAATATGAAGATCTTTATCTTTGCTTCATTTGGCTCTCATCTAGCCCA 199  
 DB 256 GAACCCAGCCCACTATGAGTTCCTTGTCTTTCATCTTGGCTCTCATGTTTCCA 315  
 QY 200 TGATTAGAGCTGATCTATCTGAGAGAAAGCTCACAGGAACGGAACATCATAGAG 259  
 DB 316 TGATTAGAGCTGATCTATCTGAGAGAAATTTTGGCTGAGATTCGGTTATG 375  
 QY 260 GATA-----TTTTCACCAATACAGCCATATCAACGATATCCACTAAATATCTCCTG 313  
 DB 376 GGTATGGCCCTTATGACCCAGTTCAGAACACCACTATACCCACCAACCAACAC 435  
 QY 314 CGT---ATCCATTTCCCTTAAATCTGCTTAGTAACCTACAGGACATGATTAGAGATT 370  
 DB 436 AATACCAACAATATACCTTTTATATCATCATGTAACCTGACGAGATGATTATTGAGGCT 495

RESULT 15  
 ADB63039  
 ID ADB63039 standard; cDNA; 1584 BP.  
 AC ADB63039;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE Human cDNA encoding clone SALGL10001070.  
 XX  
 KW Human; ss; gene; pharmaceutical; diagnostic; gene therapy;  
 KW tissue regeneration; cell regeneration; membrane protein;  
 KW signal transduction-related protein; transcription-related protein;  
 KW osteoporosis; neurological disease; cancer; tumour.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 89..997  
 FT /\*tag= a  
 FT /product= "Clone SALGL10001070 protein"  
 XX  
 EP1308459-A2.  
 XX  
 PD 07-MAY-2003.  
 XX  
 PF 28-MAR-2002; 2002EP-00007401.  
 XX  
 PR 05-NOV-2001; 2001JP-00379298.  
 PR 25-JAN-2002; 2002US-00350978.  
 XX  
 PA (HELI-) HELIX RES INST.  
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
 XX WPI; 2003-450961/43.  
 DR P-PSDB; ADB65009.  
 XX New polynucleotides and polypeptides, useful for developing a diagnostic  
 PT marker or medicines for regulation of their expression and activity, or  
 PT as targets of gene therapy.  
 XX  
 PS Claim 1; Page; 222pp; English.  
 XX The invention discloses a polynucleotide comprising a sequence selected  
 CC from 1970 fully defined nucleotide sequences which encode novel  
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide  
 CC or its partial peptide, an antibody binding to the polypeptide or peptide  
 CC of the polynucleotide, immunologically assaying the polypeptide or  
 CC peptide of the polynucleotide by contacting the polypeptide or peptide  
 CC with the antibody of the encoded protein, and observing the binding  
 CC between the two, a transformant carrying the polynucleotide in an  
 CC expressible manner and an antisense polynucleotide. The oligonucleotide  
 CC is useful as a primer for synthesising the polynucleotide, or as a probe  
 CC for detecting the polynucleotide. The polynucleotides and encoded  
 CC proteins are useful as pharmaceutical agents and many disease-related  
 CC genes may be included in them, for developing a diagnostic marker or  
 CC medicines for regulation of their expression and activity, or as targets  
 CC of gene therapy. The genes are involved in tissue and/or cell  
 CC regeneration. Membrane proteins, signal transduction-related proteins,  
 CC transcription-related proteins, disease-related proteins and genes  
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,  
 CC neurological diseases, cancer, tumours). The cDNA may be used to regulate  
 CC the activity or expression of the encoded protein to treat diseases. The  
 CC sequence presented is a cDNA of the invention. Note: Some of the sequence  
 CC data for this patent is not represented in the printed specification, but  
 CC is based on sequence information supplied by the European Patent Office.  
 XX  
 SQ Sequence 1584 BP; 526 A; 298 C; 331 G; 429 T; 0 U; 0 Other;

Query Match 11.2%; Score 67.6; DB 9; Length 1584;  
 Best Local Similarity 61.2%; Pred. No. 4.6e-05;  
 Matches 147; Conservative 0; Mismatches 84; Indels 9; Gaps 2;  
 QY 140 GGACTCCACCAATATGAAGATCTTTATCTTTGCTTCATTTGGCTCTCATCTAGCCCA 199  
 DB 1083 GAACCCAGCCCACTATGAGTTCCTTGTCTTTCATCTTGGCTCTCATGTTTCCA 1142  
 QY 200 TGATTAGAGCTGATCTATCTGAAGAGAAACGCTCACAGGAACGGAACATCATAGAG 259  
 DB 1143 TGATTGAGCTGATCTATCTGAAGAGAAATTTTGGCTGAGATTCGGTTATG 1202  
 QY 260 GATA-----TTTTCACCAATACAGCCATATCAACGATATCCACTAAATATCTCCTG 313  
 DB 1203 GGTATGGCCCTTATCAGCCAGTTCCAGAACCACTATACCCACCAACCAACAC 1262  
 QY 314 CGT---ATCCATTTCCCTTAAATGCTGCTTAGTAACCTACAGGACATGATTAGAGATT 370  
 DB 1263 AATACCAACAATATACCTTTTATATCATCATGTAACCTGACGAGACATGATTATTGAGGCT 1322

Search completed: August 15, 2004, 21:45:07  
 Job time : 355 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 21:39:53 ; Search time 73 Seconds  
(without alignments)  
4591.648 Million cell updates/sec

Title: US-10-079-754A-4

Perfect score: 604  
Sequence: 1 gaagatttcagttctata.....tatcaagcataaaaaaaa 604

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42.4	7.0	640681	4	US-09-790-988-1
C 2	42.2	7.0	640681	4	US-09-790-988-1
C 3	41.8	6.9	11131	4	US-10-204-708-28
C 4	41.4	6.9	5501	4	US-10-204-708-38
C 5	41.4	6.9	19513	4	US-10-204-708-39
C 6	41.2	6.8	1864	4	US-09-468-265-4
C 7	41	6.8	750	4	US-09-023-655-197
C 8	41	6.8	5152	4	US-10-204-708-74
C 9	41	6.8	11469	4	US-09-367-895-29
C 10	41	6.8	11478	3	US-08-981-803-29
C 11	41	6.8	11478	3	US-08-983-440-29
C 12	40.8	6.8	6152	3	US-08-973-462-1
C 13	40.4	6.7	465	6	549550-9
C 14	40.4	6.7	466	6	549550-7
C 15	40.2	6.7	6152	3	US-08-973-462-1
C 16	40	6.6	789	4	US-09-702-705-214
C 17	40	6.6	789	4	US-09-736-457-214
C 18	40	6.6	789	4	US-09-614-124B-214
C 19	40	6.6	789	4	US-09-671-325-214
C 20	40	6.6	789	4	US-09-589-184-214
C 21	40	6.6	6583	4	US-10-204-708-26
C 22	40	6.6	7664	4	US-10-204-708-84
C 23	39.8	6.6	4185	4	US-09-417-485D-7
C 24	39.8	6.6	6107	4	US-09-482-273-47
C 25	39.8	6.6	10640	4	US-09-417-485D-5
C 26	39.6	6.6	10640	4	US-09-417-485D-5
C 27	39.4	6.5	128779	4	US-09-457-855A-38

Sequence 2813, Ap  
Sequence 2813, Ap  
Sequence 13, Appl  
Sequence 1, Appl  
Sequence 13, Appl  
Sequence 2706, Ap  
Sequence 2, Appl  
Sequence 19, Appl  
Sequence 1, Appl  
Sequence 2, Appl  
Sequence 3, Appl  
Sequence 18, Appl  
Sequence 21, Appl  
Sequence 65, Appl  
Sequence 1831, Ap  
Sequence 8, Appl

US-09-790-988-1  
; Sequence 1, Application US/09790988  
; Patent No. 6632935  
; GENERAL INFORMATION:  
; APPLICANT: SHIGENOBU, SHUJI  
; APPLICANT: WATANABE, HIDEKI  
; APPLICANT: HATTORI, MASAHIRO  
; APPLICANT: SAKAKI, YOSHIYUKI  
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS  
; FILE REFERENCE: 081356/0159  
; CURRENT APPLICATION NUMBER: US/09/790,988  
; PRIOR FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: JP2000-107160  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 640681  
; TYPE: DNA  
; ORGANISM: Buchnera sp.  
US-09-790-988-1

#### ALIGNMENTS

#### RESULT 1

US-09-790-988-1  
; Sequence 1, Application US/09790988  
; Patent No. 6632935  
; GENERAL INFORMATION:  
; APPLICANT: SHIGENOBU, SHUJI  
; APPLICANT: WATANABE, HIDEKI  
; APPLICANT: HATTORI, MASAHIRO  
; APPLICANT: SAKAKI, YOSHIYUKI  
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS  
; FILE REFERENCE: 081356/0159  
; CURRENT APPLICATION NUMBER: US/09/790,988  
; PRIOR FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: JP2000-107160  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 640681  
; TYPE: DNA  
; ORGANISM: Buchnera sp.  
US-09-790-988-1

Query Match 7.0%; Score 42.4; DB 4; Length 640681;  
Best Local Similarity 47.4%; Fred. No. 1.8; Mismatches 141; Indels 0; Gaps 0;  
Matches 127; Conservative 0;

QY 39 TCTAATTACAAAAACAAATGAAGGATTTCAAGGTATTTAAACACAGCAGTCTTTCTAGCAA 98  
Db 518530 TATTAATAATTAGAAAAACAACATAAATTAATGGTTTAGAAAAAGGAAACTCCTATTAG 518589

QY 99 AGAATCTCTCTGAAGATCAGATTTTCATCTTTCATGCTGGACTCCACCAATATGAA 158  
Db 518590 ACTAATACAGAAAAAGTAGTGTAGTCTATTATTTATGATATATTAAAAAATAATGCA 518649

QY 159 GATCTTTATCTTTGCTTCATTTATCGCTCTCATCTAGCCATGATTAGAGCTGATTCATC 218  
Db 518650 AAAATTTTTCATGATTTTATAAGACAGAAAAATAAATAATTTGTTCTCCGAATTT 518709

QY 219 TGAAGAGAAACGTTCACAGAAACCGAAAAACATCATAGAGATATTTTCAACAAATACCA 278  
Db 518710 TTATATCCATCAAGATGAAGATAAAAAAGAAATATTTTGAATATTCTGTAATTTACGA 518769

QY 279 GCCATTCACAGATATCCACTAAATTTAT 306  
Db 518770 ACTATATCCAAATTTCAATATAAGAT 518797

RESULT 2  
US-09-790-988-1/c



QY 360 TAGAGAGATTTTACAAATGATTTTCCCTACTCTTCTTGTTGGTGAACCATCTTTC 419  
Db 1981 TACATTTTATTTAAAAAATAAATTTCTTCAAAACAAAATAATCTAAATATCACCTAAA 1922  
QY 420 AATGATATAAACAAGAAAAAATAATCAGTCAAGTAGTTGCAACACATATCTTGAAT 479  
Db 1921 ATAAAAATTAAACAAAAAATAATATAACAAAAAATAACAAAAAATAATATATATATAA 1862  
QY 480 CAATATCAATATTTTAAAAACATAAATAATGA 510  
Db 1861 TAAAAATATAAANAATAAATAAATAAATCA 1831

RESULT 5  
US-10-204-708-39/c  
; Sequence 39, Application US/10204708  
; Patent No. 667731  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication  
; TITLE OF INVENTION: by Assessing DNA Methylation  
; FILE REFERENCE: 5013.1012  
; CURRENT APPLICATION NUMBER: US/10/204,708  
; CURRENT FILING DATE: 2003-05-06  
; PRIOR APPLICATION NUMBER: PCT/EP01/03971  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: DE 10019058.8  
; PRIOR FILING DATE: 2000-04-06  
; PRIOR APPLICATION NUMBER: DE 10019173.8  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 98  
; SEQ ID NO 39  
; LENGTH: 19513  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-204-708-39

Query Match 6.9%; Score 41.4; DB 4; Length 19513;  
Best Local Similarity 46.9%; Pred. No. 1.2; Mismatches 146; Indels 0; Gaps 0;  
Matches 129; Conservative 0;  
QY 227 AACGTCACAGAAACGGAACAAACATCATAGAGGATATTTTCAACATACAGCCATATC 286  
Db 9628 AACTACTCAAAACTACATAAATAATCGCTTAAACCTTAAAAACAAAATACAAATA 9569  
QY 287 AACGATATCCACTAAATATCTCCTCGGTATCCATTTCTTAAATGCTGCTTAGTAAC 346  
Db 9568 AACCAAAATCGCACCCTATCTCCAACTTAAATAACAAAAACAAACCTCGTCTCAAAA 9509  
QY 347 TACAGGACATGATTAGAGAGATTTTTCACATGATTTTCTCTCTCTTCTGTTGTTG 406  
Db 9508 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 9449  
QY 407 AAAACCATCTTTCAATGAATAAACAAGAAAAAATAAATAAATAAATAAATAAATAAATAA 466  
Db 9448 AACCAACTTTTCAATAATTTCAACCAAAAAAATAAATAAATAAATAAATAAATAAATAA 9389  
QY 467 ACATACCTTGAATCAATATCAATATTTTAAACA 501  
Db 9388 ACTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 9354

RESULT 6  
US-09-468-265-4

; Sequence 4, Application US/09468265  
; Patent No. 637928  
; GENERAL INFORMATION:  
; APPLICANT: Berka, Randy M  
; APPLICANT: Cullen, Daniel  
; APPLICANT: Gray, Gregory L  
; APPLICANT: Hayenga, Kirk J  
; APPLICANT: Lawlis, Virgil B  
; TITLE OF INVENTION: Heterologous Polypeptides Expressed in Filamentous Fungi, Process  
; TITLE OF INVENTION: Making Same and Vectors for Making Same  
; FILE REFERENCE: A-42903-5  
; CURRENT APPLICATION NUMBER: US/09/468,265  
; CURRENT FILING DATE: 1999-12-10  
; PRIOR APPLICATION NUMBER: 08/484,384  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: 08/284,942  
; PRIOR FILING DATE: 1994-08-02  
; PRIOR APPLICATION NUMBER: 07/413,010  
; PRIOR FILING DATE: 1989-09-25  
; PRIOR APPLICATION NUMBER: 07/163,219  
; PRIOR FILING DATE: 1988-02-26  
; PRIOR APPLICATION NUMBER: 06/882,224  
; PRIOR FILING DATE: 1986-07-07  
; PRIOR APPLICATION NUMBER: 06/771,374  
; PRIOR FILING DATE: 1985-08-29  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 1864  
; TYPE: DNA  
; ORGANISM: Emericella nidulans  
; US-09-468-265-4  
Query Match 6.8%; Score 41.2; DB 4; Length 1864;  
Best Local Similarity 46.3%; Pred. No. 0.68;  
Matches 136; Conservative 0; Mismatches 158; Indels 0; Gaps 0;  
QY 221 AAGAGAAACGTCACAGGAAACGAAAAACATCATAGAGGATATTTTCAACATACAGC 280  
Db 1224 AAGATATCTAGCCAGGTAGACAATAATTAACCTTAAAAATATAAAAAATAGCTACTAA 1283  
QY 281 CATATCAAGGATATCCATTAATTTCTCTCGGTATCCATTTCTTAAATGCTGCTT 340  
Db 1284 AACTGAATAATTAAGAAGATAGTATTTATCTAAGTAACTAAATCTTAGTAATATA 1343  
QY 341 AGTAACACAGGACATGATTAGAGAGATTTTTCACAAATGATTTTCTCTACTCTTTCTGT 400  
Db 1344 AAGAATTTAATTAATTAATTAATAATAATAATAATAATAATAATAATAATAATAATA 1403  
QY 401 GTGTTGAAAACCATCTTTCAATGAATAAACAAGAAAAAATAAATCAGTCAAGTAGTTG 460  
Db 1404 CTTTATTATTAATTTATTAATTTATTAATTAATTAATTAATTAATTAATTAATTAAT 1463  
QY 461 CACAACACATCTTGAATCAATATCAATATTTTAAAAACATAAATAATGATAGT 514  
Db 1464 TAAAAATAGCTAGAGATTAATAAGCTCTATAAAAAATTTTAAGATTTT 1517  
RESULT 7  
US-09-023-655-197/c  
; Sequence 197, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO

```

, RESULT 8
, US-10-204-708-74/c
, ; Sequence 74, Application US/10204708
, ; Patent No. 6677731
, ; GENERAL INFORMATION:
, ; APPLICANT: OLEK, Alexander
, ; APPLICANT: PIEPENBROCK, Christian
, ; APPLICANT: BERLIN, Kurt
, ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
, ; TITLE OF INVENTION: Dy Assessing DNA Methylation
, ; FILE REFERENCE: 5013.1012
, ; CURRENT APPLICATION NUMBER: US/10/204, 708
, ; CURRENT FILING DATE: 2003-05-06
, ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
, ; PRIOR FILING DATE: 2001-04-06
, ; PRIOR APPLICATION NUMBER: DE 10019058.8
, ; PRIOR FILING DATE: 2000-04-06
, ; PRIOR APPLICATION NUMBER: DE 10019173.8

```

	Query Match	6.8%	Score 41	DB 4	Length 11469
	Best Local Similarity	51.4%	Pred. No. 1.3		
	Matches 95	Conservative	0	Mismatches 90	Indels 0
Gaps	0				
Qy	368	TTTTTCACAATGATTTTTTCTACTCTTTCTGTTGTTGAAACCATCTTTCATATGAAT	427		
Db	11372	TTTGACATCTTTTTCTCTGTATAAATTTTCTCTTTTTCATTAACCAACCAACCAATGAAG	1131		



```
QY 428 AAAACAAGAAAAAATCAAGTGTGTCACACACATCTTGGATCAATATC 487
Db 11312 AAAACAATGAAGAAATCAATAGAAAGAGGAAGATTTTCATCTCATTTGAAAGCGAA 11253
QY 488 AATATTTTAAACATATAATGATCTCTGCACTATGTAATGGTTTCTACTTTCTTT 547
Db 11252 AATCTACTAGAGAGATGATGATGAACTCTTGAACTTTTATTAGTGACACCCCTTCGTT 11193
QY 548 TCTCT 552
Db 11192 CTACT 11188

RESULT 10
US-08-981-803-29/c
; Sequence 29, Application US/08981803
; Patent No. 6147279
; GENERAL INFORMATION:
; APPLICANT: POULSEN, PETER
; TITLE OF INVENTION: INHIBITION OF GENE EXPRESSION
; FILE REFERENCE: 674509-2004
; CURRENT APPLICATION NUMBER: US/08/981,803
; CURRENT FILING DATE: 1997-04-17
; EARLIER APPLICATION NUMBER: PCT/EP96/03052
; EARLIER FILING DATE: 1996-07-12
; EARLIER APPLICATION NUMBER: 9514435.8
; EARLIER FILING DATE: 1995-07-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 11478
; TYPE: DNA
; ORGANISM: Solanum tuberosum
US-08-981-803-29

Query Match 6.8%; Score 41; DB 3; Length 11478;
Best Local Similarity 51.4%; Pred. No. 1.3;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 368 TTTTTCACATGATTTTCTCTACTCTTCTGTTGTGTTGAAACCATCTTCAATGAAT 427
Db 11381 TTGACATCTTTCTGTTATTAATTTCTTTTCAATACACACCAATGAAG 11322
QY 428 AAAACAAGAAAAAATCAAGTGTGTCACACACATCTTGGATCAATATC 487
Db 11321 AAAACAATGAAGAAATCAATAGAAAGAGGAAGATTTTCATCTCATTTGAAAGCGAA 11262
QY 488 AATATTTTAAACATATAATGATGATCTCTGCACTATGTAATGGTTTCTACTTTCTTT 547
Db 11261 AATCTACTAGAGAGATGATGATGAACTCTTGAACCTTTTCAATAGTGACACCCCTTCGTT 11202
QY 548 TCTCT 552
Db 11201 CTACT 11197

RESULT 11
US-08-983-440-29/c
; Sequence 29, Application US/08983440
; Patent No. 6232122
; GENERAL INFORMATION:
; APPLICANT: POULSEN, Peter
; TITLE OF INVENTION: INHIBITION OF GENE EXPRESSION
; FILE REFERENCE: 674509-2003
; CURRENT APPLICATION NUMBER: US/08/983,440
; CURRENT FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 9514437.4
; EARLIER FILING DATE: 1995-07-14
; EARLIER APPLICATION NUMBER: PCT/EP96/03053
; EARLIER FILING DATE: 1996-07-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0

QY 428 AAAACAAGAAAAAATCAAGTGTGTCACACACATCTTGGATCAATATC 487
Db 11321 AAAACAATGAAGAAATCAATAGAAAGAGGAAGATTTTCATCTCATTTGAAAGCGAA 11262
QY 488 AATATTTTAAACATATAATGATGATCTCTGCACTATGTAATGGTTTCTACTTTCTTT 547
Db 11261 AATCTACTAGAGAGATGATGATGAACTCTTGAACCTTTTCAATAGTGACACCCCTTCGTT 11202
QY 548 TCTCT 552
Db 11201 CTACT 11197

Query Match 6.8%; Score 40.8; DB 3; Length 6152;
Best Local Similarity 53.0%; Pred. No. 1.2;
Matches 87; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
```

```
; SEQ ID NO 29
; LENGTH: 11478
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)..(11478)
; OTHER INFORMATION: B stands for G or C or T/U
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)..(11478)
; OTHER INFORMATION: R stands for G or A
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)..(11478)
; OTHER INFORMATION: K stands for G or T/U
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)..(11478)
; OTHER INFORMATION: W stands for A or T/U
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)..(11478)
; OTHER INFORMATION: M stands for A or C
; OTHER INFORMATION: M stands for A or C
US-08-983-440-29

Query Match 6.8%; Score 41; DB 3; Length 11478;
Best Local Similarity 51.4%; Pred. No. 1.3;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 368 TTTTTCACATGATTTTCTCTACTCTTCTGTTGTGTTGAAACCATCTTCAATGAAT 427
Db 11381 TTGACATCTTTCTGTTATTAATTTCTTTTCAATACACACCAATGAAG 11322
QY 428 AAAACAAGAAAAAATCAAGTGTGTCACACACATCTTGGATCAATATC 487
Db 11321 AAAACAATGAAGAAATCAATAGAAAGAGGAAGATTTTCATCTCATTTGAAAGCGAA 11262
QY 488 AATATTTTAAACATATAATGATGATCTCTGCACTATGTAATGGTTTCTACTTTCTTT 547
Db 11261 AATCTACTAGAGAGATGATGATGAACTCTTGAACCTTTTCAATAGTGACACCCCTTCGTT 11202
QY 548 TCTCT 552
Db 11201 CTACT 11197

RESULT 12
US-08-973-462-1
; Sequence 1, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973,462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 6152
; TYPE: DNA
; ORGANISM: P. falciparum
US-08-973-462-1

Query Match 6.8%; Score 40.8; DB 3; Length 6152;
Best Local Similarity 53.0%; Pred. No. 1.2;
Matches 87; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
```

[illegible]

RESULT 13  
5496550-9  
; Patent No. 5496550  
; APPLICANT: WALLACH, MICHAEL; FUGATSCH, THERA; MENCHER, DAVID  
; TITLE OF INVENTION: METHOD OF REDUCING THE OUTPUT OF EIMERIA  
; OOCYSTS FROM A NEWBORN CHICK

	Query Match	6.7%	Score 40.4	DB 6	Length 465
	Best Local Similarity	51.7%	Pred. No. 0.71		
	Matches 92	Conservative 0	Mismatches 86	Indels 0	Gaps 0
QY	425	AATAAACAAGAAAAA	AANTCAGTCAGTAGTTC	CAACACACATCTTGGATCAAT	484
DB	149	AAGCAAGAAACACCT	CTTTATACAGATGAGTA	ATAACACATAAATAATTAATAA	208
QY	485	ATCAATATTTTAAAA	CATAATATGATAGTCT	CGAACATATGTAATTGCTTCTACTTC	544
DB	209	ATAATAATAATAATA	ATTAATTTATTTATATATTTATTTATTTATTTATTTG		268
QY	545	TTTTCTGTGCATTC	ACCATGCATCTTAAT	TAATTAATTGATCTCAAGCATAAAAAA	602
DB	269	GTGCTCTGTGATTC	GTGATGCATATTC	CACTTCAAAAAGCCGCAAAAATAA	326

RESULT 14  
5496550-7  
; Patent No. 5496550  
; APPLICANT: WALLACH, MICHAEL, PUGATSCH, THERA; MENCHER, DAVID  
; TITLE OF INVENTION: METHOD OF REDUCING THE OUTPUT OF EIMERIA  
; OOCYSTS FROM A NEWBORN CHICK  
; NUMBER OF SEQUENCES: 10  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/108,763  
; FILING DATE: 17-AUG-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 642,219  
; FILING DATE: 16-JAN-1991  
; APPLICATION NUMBER: 310,603  
; FILING DATE: 14-FEB-1989  
; APPLICATION NUMBER: 155,245  
; FILING DATE: 12-FEB-1988  
; APPLICATION NUMBER: 896,611  
; FILING DATE: 14-AUG-1986

```

;SEQ ID NO:7:
; LENGTH: 466
5496550-7

Query Match      6.7%; Score 40.4; DB 6; Length 466;
Best Local Similarity 51.7%; Fred. No. 0.71;
Matches 92; Conservative 0; Mismatches 86; Indels 0; Gaps 0

Qy  425  AATAAAACAAGAAAAAATCATGCTCAAGTAGTTCGACACACATCTTGGAAATCAAT 484
Db  150  AAGCAAGAAACACCTCTTATAGACATGATGATAAATAACACATAATAAATTAATAA 209
Qy  485  ATCAATATTTAAAAACATAAATAAGTAGTCTCTGAACATGTAATGGTTCTACTTTC 544
Db  210  ATAAATAAATAAATAATATTAATTTATTTATATTTATTTATTTATTTATTTG 269
Qy  545  TTTTCTCTGTCTACTTACCATCGATGCTTAATAATTCATCTCAAGCATAAAAAAA 602
Db  270  GGTGCTTTGTGATTCATGCTCAATGATATTCCTTCAAAAAAGCGCGAAAAATAA 327

```

```

RESULT 15
US-08-973-462-1/c
; Sequence 1, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973,462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 6152
; TYPE: DNA
; ORGANISM: P. falciparum
US-08-973-462-1

```

[illegible]

Search completed: August 16, 2004, 00:16:44  
Job time : 77 secs



QY 61 GGATTTCAAGGTATTTAAACACAGCAGTCTTTCTAGCAAAACATCTCTGAGAGCATCAG 120  
DB 61 GGATTTCAAGGTATTTAAACACAGCAGTCTTTCTAGCAAAACATCTCTGAGAGCATCAG 120  
QY 121 AATTTTCATCTTTTCATGACTGGACTCCACCAATATGAAGATCTTTATCTTTTGTCTTCATT 180  
DB 121 AATTTTCATCTTTTCATGACTGGACTCCACCAATATGAAGATCTTTATCTTTTGTCTTCATT 180  
QY 181 ATGCTCTCATCTCTAGCATGATTGAGCTGATTCATCTCTGAGAGAAACGTCACAGAA 240  
DB 181 ATGCTCTCATCTCTAGCATGATTGAGCTGATTCATCTCTGAGAGAAACGTCACAGAA 240  
QY 241 CGGAAAAAATCATATAGAGGATATTTTCAACAATACCAGCCATATCAACGATATCCACTA 300  
DB 241 CGGAAAAAATCATATAGAGGATATTTTCAACAATACCAGCCATATCAACGATATCCACTA 300  
QY 301 AATTATCTCTCTGATTCATCTTTTCTTAAATGCTGCTTAGTAATCAACAGGATGATT 360  
DB 301 AATTATCTCTCTGATTCATCTTTTCTTAAATGCTGCTTAGTAATCAACAGGATGATT 360  
QY 361 AGAGAGATTTTTCACATGATTTTCTCTACTCTTTCTGTTGTTGAAACCATCTTTCA 420  
DB 361 AGAGAGATTTTTCACATGATTTTCTCTACTCTTTCTGTTGTTGAAACCATCTTTCA 420  
QY 421 AATGAATATAAACAAGAAAAAATCAAGTCAAGTGTGCAACAACATCTTGGAAATC 480  
DB 421 AATGAATATAAACAAGAAAAAATCAAGTCAAGTGTGCAACAACATCTTGGAAATC 480  
QY 481 AATATCAATATTTTAAACATATATGATGCTCTGAACTATGTAATGGTTTCTAC 540  
DB 481 AATATCAATATTTTAAACATATATGATGCTCTGAACTATGTAATGGTTTCTAC 540  
QY 541 TTTCTTTTCTCTGCTACCTACCATGCTGCTTAAATATGATCTATCAAGCATAAAAAA 600  
DB 541 TTTCTTTTCTCTGCTACCTACCATGCTGCTTAAATATGATCTATCAAGCATAAAAAA 600  
QY 601 AAAA 604  
DB 601 AAAA 604

## RESULT 2

US-10-079-754A-1  
; Sequence 1, Application US/10079754A  
; Publication No. US20020164625A1  
; GENERAL INFORMATION:  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Molenaar, Adrian J.  
; APPLICANT: Davis, Stephen R.  
; TITLE OF INVENTION: Compositions Isolated from Bovine  
; TITLE OF INVENTION: Mammary Gland and Methods for Their Use  
; FILE REFERENCE: 11000.1068  
; CURRENT APPLICATION NUMBER: US/10/079,754A  
; PRIOR FILING DATE: 2002-02-19  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 505  
; TYPE: DNA  
; ORGANISM: Bovine  
US-10-079-754A-1

Query Match

77.1%; Score 465.4; DB 14; Length 505;

Best Local Similarity 99.8%; Pred. No. 2.1e-96;  
Matches 466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 138 CTGACTCCCAAAATATGAAGATCTTTATCTTTGTCTTCTATATGCTCTCATCTAGC 197  
DB 28 CAGGACTCCCAAAATATGAAGATCTTTATCTTTGTCTTCTATATGCTCTCATCTAGC 87  
QY 198 CATGATTAGAGCTGATTCATCTGAGAGAAACGTCACAGGAAACGAAAAACATCATAG 257  
DB 88 CATGATTAGAGCTGATTCATCTGAGAGAAACGTCACAGGAAACGAAAAACATCATAG 147  
QY 258 AGGATATTTTCAACAATACCAGCCATATCAACGATATCCACTAATATCTCTCTGCTA 317  
DB 148 AGGATATTTTCAACAATACCAGCCATATCAACGATATCCACTAATATCTCTCTGCTA 207  
QY 318 TCCATTTCTTTAAATGCTGCTTAGTAATACAGGACATGATTAGAGAGATTTTTCACAA 377  
DB 208 TCCATTTCTTTAAATGCTGCTTAGTAATACAGGACATGATTAGAGAGATTTTTCACAA 267  
QY 378 TGATTTTCTTACTCTTTCTGTTGTTGAAACCATCTTTTCAAAATGAATAAACAAGA 437  
DB 268 TGATTTTCTTACTCTTTCTGTTGTTGAAACCATCTTTTCAAAATGAATAAACAAGA 327  
QY 438 AAAAAAATCAGTCAAGTGTGCAACAACATCTTGGAAATCAAAATATCAATATTTTAA 497  
DB 328 AAAAAAATCAGTCAAGTGTGCAACAACATCTTGGAAATCAAAATATCAATATTTTAA 387  
QY 498 AACATAAATATGATGCTCTGAACTATGTAATGTTGTTTCTTCTTCTCTGTCAC 557  
DB 388 AACATAAATATGATGCTCTGAACTATGTAATGTTGTTTCTTCTTCTCTGTCAC 447  
QY 558 TTACATGCTGCTTAAATATGATCTATCAAGCATAAAAA 604  
DB 448 TTACATGCTGCTTAAATATGATCTATCAAGCATAAAAA 494

## RESULT 3

US-10-079-754A-2  
; Sequence 2, Application US/10079754A  
; Publication No. US20020164625A1  
; GENERAL INFORMATION:  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Molenaar, Adrian J.  
; APPLICANT: Davis, Stephen R.  
; TITLE OF INVENTION: Compositions Isolated from Bovine  
; TITLE OF INVENTION: Mammary Gland and Methods for Their Use  
; FILE REFERENCE: 11000.1068  
; CURRENT APPLICATION NUMBER: US/10/079,754A  
; PRIOR FILING DATE: 2002-02-19  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 585  
; TYPE: DNA  
; ORGANISM: Bovine  
US-10-079-754A-2

Query Match

74.2%; Score 448.2; DB 14; Length 585;

Best Local Similarity 99.3%; Pred. No. 1.9e-92;  
Matches 450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 138 CTGACTCCCAAAATATGAAGATCTTTATCTTTGTCTTCTATATGCTCTCATCTAGC 197  
DB 133 CAGGACTCCCAAAATATGAAGATCTTTATCTTTTCTTCTTCTTCTTCTCTCTAGC 192

Qy	198	CATGATTAGAGCTGATTCAATCTGAAGAGAAACGTCACAGGAAACGGAAAAACATCATG	257
Db	193	CATGATTAGAGCTGATTCAATCTGAAGAGAAACGTCACAGGAAACGGAAAAACATCATG	252
Qy	258	AGGATATTTTCAACAATACCGCCATATCAACGATATCCCACTAAATATTCCTCTCGCGTA	317
Db	253	AGGATATTTTCAACAATACCGCCATATCAACGATATCCCACTAAATATTCCTCTCGCGTA	312
Qy	318	TCCATTTTCCTTAAAAATGCTGTTAGTAACACAGGACATGATTAGAGAGATTTTTCACAA	377
Db	313	TCCATTTTCCTTAAAAATGCTGTTAGTAACACAGGACATGATTAGAGAGATTTTTCACAA	372
Qy	378	TGATTTTTCCTACTCTTCTGTTGTTGAAAAACATCTTTCAAATGAATAAACAAGA	437
Db	373	TGATTTTTCCTACTCTTCTGTTGTTGAAAAACATCTTTCAAATGAATAAACAAGA	432
Qy	438	AAAAAAATCAGTCAAGTAGTTGCACAACACATACTTGGAAATCAATATCAATATTTTAA	497
Db	433	AAAAAAATCAGTCAAGTAGTTGCACAACACATACTTGGAAATCAATATCAATATTTTAA	492
Qy	498	AACATAAATAATGATAGTCTCTGAACTAGTAAATGGTTTCTACTTTCTTTTCTCTGTCAC	557
Db	493	AACATAAATAATGATAGTCTCTGAACTAGTAAATGGTTTCTACTTTCTTTTCTCTGTCAC	552
Qy	558	TTACCATGCATGCTTAATAAATGATCTATCAA	590
Db	553	TTACCATGCATGCTTAATAAATGATCTATCAA	585

```

RESULT 4
US-10-079-754A-6
; Sequence 6, Application US/10079754A
; Publication No. US2002016425A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; TITLE OF INVENTION: Mammary Gland and Methods for Their Use
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Bovine
; US-10-079-754A-6

```

Query Match	70.11%	Score 423.4	DB 14	Length 525
Best Local Similarity	93.4%	Pred. NO. 8.6e-87		
Matches 466	Conservative	Mismatches 0	Indels 32	Gaps 1
QY	138	CTGGACTCCACCAATATGAAGATCTTTATCTTTGTCTCTATTATGGCTCTCATCCTTAGC	197	
Db	27	CAGGATCCACCAATATGAAGATCTTTATCTTTGTCTCTATTATGGCTCTCATCCTTAGC	86	
QY	198	CATGATTAGCTGATTTCATCTGAAGAGAAACGTCACAGGAAACGGAAAAACATCAT--	255	
Db	87	CATGATTAGCTGATTTCATCTGAAGAGAAACGTCACAGGAAACGGAAAAACATCATGT	146	
QY	256	-----AGAGGATATTTTCAACAATACAGGCATAT	285	

147	TGATAGGCTCCAGAA	TCTTACTACTACAAGAGGATATTTTCAACAATACCAGCCATAT	206
286	CAACGATATCCACTAAAT	TATCCTCTGGGTATCCATTTCCITTAATAATGCTGTAGTAA	345
207	CAACGATATCCACTAAAT	TATCCTCTGGGTATCCATTTCCITTAATAATGCTGTAGTAA	266
346	CTACAGGACATGATTA	TAGAGAGATTTTTTCACAATGATTTTTTCTACTCTTTCTGTGTGT	408
267	CTACAGGACATGATTA	TAGAGAGATTTTTTCACAATGATTTTTTCTACTCTTTCTGTGTGT	326
406	GAATAACATCTTTTCA	ATAGATTAATAACAAGAAAAAATAATCAGTCAAGTAGTTGCACAA	465
327	GAATAACATCTTTTCA	ATAGATTAATAACAAGAAAAAATAATCAGTCAAGTAGTTGCACAA	386
466	CACATATCTTGGAAAT	CAAAATCAATATTTTAAACACATAATAATGATGCTCTCGAATCT	525
387	CACATATCTTGGAAAT	CAAAATCAATATTTTAAACACATAATAATGATGCTCTCGAATCT	446
526	GTAATTGGTTTCTACT	TTTCTCTGTGCACATTACCATGCATGCTCTTAATAAATTGATCT	585
447	GTAATTGGTTTCTACT	TTTCTCTGTGCACATTACCATGCATGCTCTTAATAAATTGATCT	506
586	ATCAAGCATAAAAA	AAAAA 604	
507	ATCAAGCATAAAAA	AAAAA 525	

```

RESULT 5
US-10-079-754A-15
; Sequence 15, Application US/10079754A
; Publication No. US20020164625A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; APPLICANT: Davis, Stephen R.
; TITLE OF INVENTION: Compositions Isolated from Bovine
; TITLE OF INVENTION: Mammary Gland and Methods for Their Use
; FILE REFERENCE: 11000.1068
; CURRENT APPLICATION NUMBER: US/10/079,754A
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 09/699,146
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60,162,701
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 09/644,190
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60,150,330
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Bovine
US-10-079-754A-15

```

	Query Match	35.5%;	Score 238.4;	DB 14;	Length 267;
	Best Local Similarity	99.6%;	Pred. No. 1.2e-44;		
	Matches 239;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	138	CTGACATCCACCAATATGAGATCTTTATCTTTGTCTTCATTATGGCTCTCATCTCTGAGC	197		
DB	28	CAGGACATCCACCAATATGAGATCTTTATCTTTGTCTTCATTATGGCTCTCATCTCTGAGC	87		
QY	198	CATGATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGGAAACGGAAAAAATCATCATAG	257		
DB	88	CATGATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGGAAACGGAAAAAATCATCATAG	147		
QY	258	AGGATATTTTCAACAAATACAGGCATATCAACGATATCCACTAAATATTCCTCTCGGTA	317		
DB	148	AGGATATTTTCAACAAATACAGGCATATCAACGATATCCACTAAATATTCCTCTCGGTA	207		

QY 318 TCACATTTCCCTAAATGCTCTAGTAAGTACACAGGACATGATTAGAGATTTTTCACAA 377  
 Db 208 TCACATTTCCCTAAATGCTCTAGTAAGTACACAGGACATGATTAGAGATTTTTCACAA 267

RESULT 6  
 US-10-079-623-200  
 ; Sequence 200, Application US/10079623  
 ; Publication No. US20020169302A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Havukkala, Ilkka J.  
 ; APPLICANT: Glenn, Matthew  
 ; APPLICANT: Grigor, Murray R.  
 ; APPLICANT: Molenaar, Adrian J.  
 ; TITLE OF INVENTION: Compositions isolated from bovine  
 ; TITLE OF INVENTION: mammary gland and methods for their use.

FILE REFERENCE: 11000.1044c3  
 CURRENT FILING DATE: 2002-02-19  
 NUMBER OF SEQ ID NOS: 370  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO 200  
 LENGTH: 267  
 TYPE: DNA  
 ORGANISM: Bovine  
 US-10-079-623-200

Query Match 39.5%; Score 238.4; DB 14; Length 267;  
 Best Local Similarity 99.6%; Pred. No. 1.2e-44; Indels 0; Gaps 0;  
 Matches 239; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 138 CTGACTCCCAAAATGATGAAGATCTTTATCTTTGTCTTCATATGCTCTCATCTAGC 197  
 Db 28 CAGGACTCCCAAAATGATGAAGATCTTTATCTTTGTCTTCATATGCTCTCATCTAGC 87  
 QY 198 CATGATTAGAGCTGATTCATCTGAAGAGAACTCTACAGGAAACGGAAAAACATCATAG 257  
 Db 88 CATGATTAGAGCTGATTCATCTGAAGAGAACTCTACAGGAAACGGAAAAACATCATAG 147  
 QY 258 AGGATATTTTCAACAATPACAGGCATATCAACGATATCCACTAAATATCTCTCGTA 317  
 Db 148 AGGATATTTTCAACAATPACAGGCATATCAACGATATCCACTAAATATCTCTCGTA 207  
 QY 318 TCACATTTCCCTAAATGCTCTAGTAAGTACACAGGACATGATTAGAGATTTTTCACAA 377  
 Db 208 TCACATTTCCCTAAATGCTCTAGTAAGTACACAGGACATGATTAGAGATTTTTCACAA 267

RESULT 7  
 US-10-079-754A-3  
 ; Sequence 3, Application US/10079754A  
 ; Publication No. US20020164625A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Glenn, Matthew  
 ; APPLICANT: Grigor, Murray R.  
 ; APPLICANT: Molenaar, Adrian J.  
 ; APPLICANT: Davis, Stephen R.  
 ; TITLE OF INVENTION: Compositions Isolated from Bovine  
 ; TITLE OF INVENTION: Mammary Gland and Methods for Their Use  
 ; FILE REFERENCE: 11000.1068  
 CURRENT FILING DATE: 2002-02-19  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 3  
 ; LENGTH: 869  
 ; TYPE: DNA  
 ; ORGANISM: Bovine  
 US-10-079-754A-3

Query Match 19.0%; Score 115; DB 14; Length 869;  
 Best Local Similarity 92.4%; Pred. No. 3.3e-16;  
 Matches 121; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 138 CTGACTCCCAAAATGATGAAGATCTTTATCTTTGTCTTCATATGCTCTCATCTAGC 197  
 Db 56 CAGGACTCCCAAAATGATGAAGATCTTTATCTTTGTCTTCATATGCTCTCATCTAGC 115  
 QY 198 CATGATTAGAGCTGATTCATCTGAAGAGAACTCTACAGGAAACGGAAAAACATCATAG 257  
 Db 116 CATGATTAGAGCTGATTCATCTGAAGAGAACTCTACAGGAAACGGAAAAACATCATAG 175  
 QY 258 AGGATATTTTTC 268  
 Db 176 ATGATTCTTC 186

RESULT 8  
 US-10-079-754A-5  
 ; Sequence 5, Application US/10079754A  
 ; Publication No. US20020164625A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Glenn, Matthew  
 ; APPLICANT: Grigor, Murray R.  
 ; APPLICANT: Molenaar, Adrian J.  
 ; APPLICANT: Davis, Stephen R.  
 ; TITLE OF INVENTION: Compositions Isolated from Bovine  
 ; TITLE OF INVENTION: Mammary Gland and Methods for Their Use  
 ; FILE REFERENCE: 11000.1068  
 CURRENT FILING DATE: 2002-02-19  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 5  
 ; LENGTH: 96  
 ; TYPE: DNA  
 ; ORGANISM: Bovine  
 US-10-079-754A-5

Query Match 14.4%; Score 87; DB 14; Length 96;  
 Best Local Similarity 94.7%; Pred. No. 2.8e-10;  
 Matches 90; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 265 TTTCAACAATACAGGCATATCAACGATATCCACTAAATATCTCTCGTATCCATTT 324  
 Db 1 TTTCAACAATACAGGCATATCAACGATATCCACTAAATATCTCTCGTATCCATTT 60  
 QY 325 CCTTAAATGCTGCTTAGTAAGTACACAGGACATGAT 359  
 Db 61 TCTTAAATGCTGCTTAGTAAGTACACAGGACATGAT 95

RESULT 9  
 US-09-992-600A-5  
 ; Sequence 5, Application US/09992600A  
 ; Publication No. US2003002716A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Benjanin, Stephane  
 ; APPLICANT: Tanaka, Hiroaki

;; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
;; FILE REFERENCE: 91.US4.DIV  
;; CURRENT APPLICATION NUMBER: US/09/992.600A  
;; CURRENT FILING DATE: 2001-11-13  
;; PRIOR APPLICATION NUMBER: US 09/924,340  
;; PRIOR FILING DATE: 2001-08-06  
;; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
;; PRIOR FILING DATE: 2001-08-06  
;; PRIOR APPLICATION NUMBER: US 60/305,456  
;; PRIOR FILING DATE: 2001-07-13  
;; PRIOR APPLICATION NUMBER: US 60/302,277  
;; PRIOR FILING DATE: 2001-06-29  
;; PRIOR APPLICATION NUMBER: US 60/298,698  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: US 60/293,574  
;; PRIOR FILING DATE: 2001-05-25  
;; NUMBER OF SEQ ID NOS: 114  
;; SOFTWARE: JPatent  
;; SEQ ID NO 5  
;; LENGTH: 438  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: 5'UTR  
;; LOCATION: 1..83  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 84..317  
;; FEATURE:  
;; NAME/KEY: polyA signal  
;; LOCATION: 318..438  
;; FEATURE:  
;; NAME/KEY: polyA site  
;; LOCATION: 423..438  
;; SOFTWARE: JPatent  
;; SEQ ID NO 5  
;; LENGTH: 438  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: 5'UTR  
;; LOCATION: 1..83  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 84..317  
;; FEATURE:  
;; NAME/KEY: polyA signal  
;; LOCATION: 318..438  
;; FEATURE:  
;; NAME/KEY: polyA site  
;; LOCATION: 423..438  
US-09-992-600A-5

Query Match 11.7%; Score 70.4; DB 10; Length 438;  
Best Local Similarity 69.9%; Pred. No. 3.6e-06;  
Matches 95; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 120 GAATTTCATCTTTTCATGACTGGACTCCACCAAAATATGAAGATCTTTATCTTTGCTTCAT 179  
DB 50 GACTCTCTCTTTGAGTAAAGGACTCAGCCAACTATGAAGTTTGTGCTTTGCTTTAGT 109

QY 180 TATGGCTCTCATCTAGCCATGATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGGAA 239  
DB 110 CTGGCTCTCATGATTTCCATGATTAGCGCTGATTCACATGAAAGAGACATCATGGGTA 169

QY 240 ACGGAAAAAACATCAT 255  
DB 170 TAGAGAAAATTCAT 185

RESULT 10  
US-09-924-340-5  
; Sequence 5, Application US/09924340  
; Publication No. US20030027248A1  
; GENERAL INFORMATION:  
; APPLICANT: Bejanin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/924,340  
; CURRENT FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: JPatent  
; SEQ ID NO 5  
; LENGTH: 438  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: 1..83  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 84..317  
; FEATURE:  
; NAME/KEY: polyA signal  
; LOCATION: 318..438  
; FEATURE:  
; NAME/KEY: polyA site  
; LOCATION: 423..438

;; PRIOR APPLICATION NUMBER: US 60/293,574  
;; PRIOR FILING DATE: 2001-05-25  
;; NUMBER OF SEQ ID NOS: 112  
;; SOFTWARE: JPatent  
;; SEQ ID NO 5  
;; LENGTH: 438  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: 5'UTR  
;; LOCATION: 1..83  
;; NAME/KEY: CDS  
;; LOCATION: 84..317  
;; NAME/KEY: 3'UTR  
;; LOCATION: 318..438  
;; NAME/KEY: polyA signal  
;; LOCATION: 397..402  
;; NAME/KEY: polyA site  
;; LOCATION: 423..438  
US-09-924-340-5

Query Match 11.7%; Score 70.4; DB 10; Length 438;  
Best Local Similarity 69.9%; Pred. No. 3.6e-06;  
Matches 95; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 120 GAATTTCATCTTTTCATGACTGGACTCCACCAAAATATGAAGATCTTTATCTTTGCTTCAT 179  
DB 50 GACTCTCTCTTTGAGTAAAGGACTCAGCCAACTATGAAGTTTGTGCTTTGCTTTAGT 109

QY 180 TATGGCTCTCATCTAGCCATGATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGGAA 239  
DB 110 CTGGCTCTCATGATTTCCATGATTAGCGCTGATTCACATGAAAGAGACATCATGGGTA 169

QY 240 ACGGAAAAAACATCAT 255  
DB 170 TAGAGAAAATTCAT 185

RESULT 11  
US-09-992-095B-5  
; Sequence 5, Application US/0992095B  
; Publication No. US20030157485A1  
; GENERAL INFORMATION:  
; APPLICANT: Bejanin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US5.DIV  
; CURRENT APPLICATION NUMBER: US/09/992,095B  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPatent  
; SEQ ID NO 5  
; LENGTH: 438  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: 1..83  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 84..317  
; FEATURE:  
; NAME/KEY: polyA signal  
; LOCATION: 318..438  
; FEATURE:  
; NAME/KEY: polyA site  
; LOCATION: 423..438

FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 318..438  
FEATURE:  
NAME/KEY: polyA signal  
LOCATION: 397..402  
FEATURE:  
NAME/KEY: polyA site  
LOCATION: 423..438  
US-09-992-095B-5

Query Match 11.7%; Score 70.4; DB 10; Length 438;  
Best Local Similarity 69.9%; Pred. No. 3.6e-06;  
Matches 95; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 120 GAATTTTCATCTTTCATGACTGGACTCCACCAAAATATGAAGATCTTTATCTTTGCTTCAT 179  
Db 50 GACTCTCTCTTGTAGTAAAGGACTCAGCCAACTATGAAGTTTTTTTGTCTTTGCTTAGT 109

QY 180 TATGGCTCTCATCTAGCCATGATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGAA 239  
Db 110 CTGGCTCTCATGATTTCCATGATTAGCGCTGATTCACATGAAAGAGACATCATGGTA 169

QY 240 ACGGAAAAAACATCAT 255  
Db 170 TAGAAGAAAAATTCAT 185

RESULT 12  
US-09-999-570-5  
Sequence 5, Application US/09999570  
Publication No. US20030170628A1  
GENERAL INFORMATION:  
APPLICANT: Benjamin, Stephanie  
APPLICANT: Tanaka, Hiroaki  
TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
FILE REFERENCE: G-091US08DIV  
CURRENT APPLICATION NUMBER: US/09/999,570  
CURRENT FILING DATE: 2001-06-14  
PRIOR APPLICATION NUMBER: US 09/924,340  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: PCT/IB01/01715  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: US 60/305,456  
PRIOR FILING DATE: 2001-07-13  
PRIOR APPLICATION NUMBER: US 60/302,277  
PRIOR FILING DATE: 2001-07-13  
PRIOR APPLICATION NUMBER: US 60/298,698  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 60/293,574  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 112  
SOFTWARE: JPatent  
SEQ ID NO 5  
LENGTH: 438  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..83  
NAME/KEY: CDS  
LOCATION: 84..317  
NAME/KEY: 3'UTR  
LOCATION: 318..438  
NAME/KEY: polyA signal  
LOCATION: 397..402  
NAME/KEY: polyA site  
LOCATION: 423..438  
US-09-999-570-5

US-09-999-570-5

US-10-000-489-5  
Sequence 5, Application US/10000489  
Publication No. US20030092011A1  
GENERAL INFORMATION:  
APPLICANT: Benjamin, Stephanie  
APPLICANT: Tanaka, Hiroaki  
TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
FILE REFERENCE: 91.US6.DIV  
CURRENT APPLICATION NUMBER: US/10/000,489  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: US 09/924,340  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: PCT/IB01/01715  
PRIOR FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: US 60/305,456  
PRIOR FILING DATE: 2001-07-13  
PRIOR APPLICATION NUMBER: US 60/302,277  
PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: US 60/298,698  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 60/293,574  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 112  
SOFTWARE: JPatent  
SEQ ID NO 5  
LENGTH: 438  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..83  
NAME/KEY: CDS  
LOCATION: 84..317  
NAME/KEY: 3'UTR  
LOCATION: 318..438  
NAME/KEY: polyA signal  
LOCATION: 397..402  
NAME/KEY: polyA site  
LOCATION: 423..438  
US-10-000-489-5

Query Match 11.7%; Score 70.4; DB 15; Length 438;  
Best Local Similarity 69.9%; Pred. No. 3.6e-06;  
Matches 95; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 120 GAATTTTCATCTTTCATGACTGGACTCCACCAAAATATGAAGATCTTTATCTTTGCTTCAT 179  
Db 50 GACTCTCTCTTGTAGTAAAGGACTCAGCCAACTATGAAGTTTTTTTGTCTTTGCTTAGT 109

QY 180 TATGGCTCTCATCTAGCCATGATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGAA 239  
Db 110 CTGGCTCTCATGATTTCCATGATTAGCGCTGATTCACATGAAAGAGACATCATGGTA 169

QY 240 ACGGAAAAAACATCAT 255

Query Match 11.7%; Score 70.4; DB 15; Length 438;  
Best Local Similarity 69.9%; Pred. No. 3.6e-06;  
Matches 95; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 120 GAATTTTCATCTTTCATGACTGGACTCCACCAAAATATGAAGATCTTTATCTTTGCTTCAT 179  
Db 50 GACTCTCTCTTGTAGTAAAGGACTCAGCCAACTATGAAGTTTTTTTGTCTTTGCTTAGT 109

QY 180 TATGGCTCTCATCTAGCCATGATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGAA 239  
Db 110 CTGGCTCTCATGATTTCCATGATTAGCGCTGATTCACATGAAAGAGACATCATGGTA 169

QY 240 ACGGAAAAAACATCAT 255



Search completed: August 16, 2004, 00:23:18  
Job time : 391 secs

RESULT 15  
US-10-154-678-5  
; Sequence 5, Application US/10154678  
; Publication No. US20030162186A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjanin, Stephane  
; APPLICANT: Iwanka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 182.US1.REG  
; CURRENT APPLICATION NUMBER: US/10/154.678



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 07:37:23 ; Search time 2272 Seconds  
(without alignments)  
7938.713 Million cell updates/sec

Title: US-10-079-754A-4  
Perfect score: 604  
Sequence: 1 gaagatttttcagttctata.....tatcaagcataaaaaaaa 604

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST:\*
- 1: em\_estba:\*
  - 2: em\_esthum:\*
  - 3: em\_estin:\*
  - 4: em\_estnu:\*
  - 5: em\_estov:\*
  - 6: em\_estpl:\*
  - 7: em\_estro:\*
  - 8: em\_hic:\*
  - 9: gb\_est1:\*
  - 10: gb\_est2:\*
  - 11: gb\_hic:\*
  - 12: gb\_est3:\*
  - 13: gb\_est4:\*
  - 14: gb\_est5:\*
  - 15: em\_estfun:\*
  - 16: em\_estom:\*
  - 17: em\_gss\_hum:\*
  - 18: em\_gss\_inv:\*
  - 19: em\_gss\_pln:\*
  - 20: em\_gss\_vrt:\*
  - 21: em\_gss\_fun:\*
  - 22: em\_gss\_mam:\*
  - 23: em\_gss\_mus:\*
  - 24: em\_gss\_pro:\*
  - 25: em\_gss\_rod:\*
  - 26: em\_gss\_phg:\*
  - 27: em\_gss\_vrl:\*
  - 28: gb\_gss1:\*
  - 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82.2	13.6	553	14	CB985810
2	78.6	13.0	546	14	CB957544
3	78.6	13.0	555	14	CB986888
4	78.6	13.0	792	14	CB958289

5	78.6	13.0	793	14	CB984475
6	78.6	13.0	831	14	CB985521
7	78.6	13.0	929	14	CB987745
8	77.6	12.8	583	13	EX485825
9	77.6	12.8	622	13	EX485714
10	77.6	12.8	773	14	CB987175
11	77	12.7	312	9	AA376700
12	77	12.7	327	9	AA382524
13	77	12.7	330	9	AA376704
14	77	12.7	348	9	AA376689
15	77	12.7	390	13	EX283872
16	77	12.7	527	14	CB986478
17	77	12.7	533	14	CB985159
18	77	12.7	540	14	CB957853
19	77	12.7	543	14	CB958523
20	77	12.7	543	14	CB987834
21	77	12.7	544	14	CB957223
22	77	12.7	544	14	CB959166
23	77	12.7	545	14	CB958545
24	77	12.7	545	14	CB959029
25	77	12.7	546	14	CB956307
26	77	12.7	546	14	CB956449
27	77	12.7	546	14	CB957121
28	77	12.7	546	14	CB957972
29	77	12.7	546	14	CB985184
30	77	12.7	546	14	CB986281
31	77	12.7	546	14	CB986417
32	77	12.7	546	14	CB987123
33	77	12.7	546	14	CB987595
34	77	12.7	547	14	CB955720
35	77	12.7	547	14	CB956188
36	77	12.7	547	14	CB956487
37	77	12.7	547	14	CB956536
38	77	12.7	547	14	CB956860
39	77	12.7	547	14	CB956877
40	77	12.7	547	14	CB957224
41	77	12.7	547	14	CB957280
42	77	12.7	547	14	CB957414
43	77	12.7	547	14	CB957440
44	77	12.7	547	14	CB957523
45	77	12.7	547	14	CB957940

ALIGNMENTS

RESULT 1  
CB985810  
LOCUS  
DEFINITION  
IMAGE:30329983 5', mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

CB985810 553 bp mRNA linear EST 01-MAY-2003  
AGENCOURT 13647469 NIH MGC 184 Homo sapiens cDNA clone  
IMAGE:30329983 5', mRNA sequence.

CB985810  
CB985810.1 GI:30280334  
EST.  
Homo sapiens (human)

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC <http://mgs.nci.nih.gov/>  
1 (bases 1 to 553)

National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits

cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: NDCM143 row: b column: 08  
High quality sequence stop: 553.

FEATURES  
source

Location/Qualifiers  
1. .553  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30329983"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_184"  
/notes="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site\_1: Sfil (ggccatagggc); Site\_2: Sfil (ggcgctcgcc); Library is oligo-dr primed and directionally cloned. cDNA was prepared from a glandular pool of tissues from thyroid, parathyroid, adrenal, cortex and pineal gland. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGAGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.38 Kb (range 0.60-3.5 Kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 13.6%; Score 82.2; DB 14; Length 553;  
Best Local Similarity 62.8%; Pred. No. 0.00043;  
Matches 145; Conservative 0; Mismatches 83; Indels 3; Gaps 1;  
QY 140 GGACTCCACCAATATGAAGATCTTTATCTTTGCTTCATTATGGCTCTCATCTAGCCA 199  
DB 62 GAACCCACCAACTATGAAGTTCCTCTCTTGGCTCTCATCTGGCTCTCATGTTTCA 121  
QY 200 TGATTAGAGCTGATTCATCTGAAGAGAAAGCTCAGAGAAACGGAACAAACATCATAGAG 259  
DB 122 TGATTGAGCTGATTCATCTGAAGAGAAATTTTGGCTAGAAATGGAAGATTCGGTTATG 181  
QY 260 GATATTTTCAACATACACGACCATATCAAGCATATCCACTAAATTTATCTCTCGGTATC 319  
DB 182 GGTA---TGGCCCTATACGCCAGTTCAGAAACACCATATACCCACCAACATACCAAC 238  
QY 320 CATTTCTTAAATGCTGCTTAGTACTACAGACATGATTAGAGATTT 370  
DB 239 AATATACCTTTAATATCATCATGTAATCGAGGACATGATTATTAGAGCTT 289

RESULT 2  
CB986888

LOCUS CB986888 546 bp mRNA linear EST 29-APR-2003  
DEFINITION AGENCOURT\_13785762 NIH\_MGC\_184 Homo sapiens cDNA clone  
IMAGE:30352504 5', mRNA sequence.  
ACCESSION CB986888  
VERSION CB986888.1 GI:30213661  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 546)  
TITLE NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbe-remail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
plate: NDCM151 row: 1 column: 17  
High quality sequence stop: 546.

FEATURES  
source

Location/Qualifiers  
1. .546

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30352504"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_184"  
/notes="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site\_1: Sfil (ggccatagggc); Site\_2: Sfil (ggcgctcgcc); Library is oligo-dr primed and directionally cloned. cDNA was prepared from a glandular pool of tissues from thyroid, parathyroid, adrenal, cortex and pineal gland. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGAGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.38 Kb (range 0.60-3.5 Kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 13.0%; Score 78.6; DB 14; Length 546;  
Best Local Similarity 60.2%; Pred. No. 0.0017;  
Matches 151; Conservative 0; Mismatches 94; Indels 6; Gaps 1;  
QY 120 GAATTTTCATCTTTTCATGACTGAGCTCCACCAATATGAAGATCTTTATCTTTGTTTCAT 179  
DB 41 GACTCTCTCTTCTGAGTAAAGGACTCAGCCAACTATGAAGTCTTTGTCTTTGCTTTAGT 100  
QY 180 TATGGCTTCATCTAGCCATGATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGGAA 239  
DB 101 CTGGCTCTCATGATTTCCATGATTAGCGCTGATTCATGATGAAGAGACATCATGGTA 160  
QY 240 ACGAAACAAACATCATAGAGGATATTTTCAACATACCGCCATATCAACGATATCCACT 299  
DB 161 TAGAAGAAATTCAT-----GAAAGCATCATTCACATCGAATTTCCATTTATGG 214  
QY 300 AATATCTCTCGCGCTATCTCCATTTCTTAAATGCTGCTTAGTAACATACAGGACATGAT 359  
DB 215 GGACTATGATCAATATCTATATGACAAATGATATCTTAGTATCATCGGGCATGAT 274  
QY 360 TAGAGAGATTT 370  
DB 275 TATAGAGGTTT 285

RESULT 3  
CB986888

LOCUS CB986888 555 bp mRNA linear EST 01-MAY-2003  
DEFINITION AGENCOURT\_13650592 NIH\_MGC\_184 Homo sapiens cDNA clone  
IMAGE:30329446 5', mRNA sequence.  
ACCESSION CB986888  
VERSION CB986888.1 GI:30281408  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 555)  
TITLE NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
plate: NDCM141 row: k column: 23  
High quality sequence stop: 551.

## FEATURES

source Location/Qualifiers  
 1. .555  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="IMAGE:30329446"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC 184"  
 /notes="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site\_1: SfiI (ggccattggcc); Site 2: SfiI (ggccgctggcc); Library is oligo-dr primed and directionally cloned. cDNA was prepared from a glandular pool of tissues from thyroid, parathyroid, adrenal, cortex and pineal gland. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.38 kb (range 0.60-3.5 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 13.0%; Score 78.6; DB 14; Length 555;  
 Best Local Similarity 60.2%; Pred. No. 0.0016;  
 Matches 151; Conservative 0; Mismatches 94; Indels 6; Gaps 1;  
 QY 120 GAATTTTCATCTTTTCATGACTGGACTCCACCATAATATGAAGATCTTTATCTTTGTCTTCAT 179  
 Db 41 GACTCTCTCTTGGTAAAGGACTCAGCCAACTATGAAGTTTITGCTTTGCTTTAGT 100  
 QY 180 TATGGCTCTCATCCTAGCCTAGCTGATTTAGACTGATTCATCTGAAGAGAACTGACAGGAA 239  
 Db 101 CTTCGCTCTCATGATTTTCATGATTTAGCGCTGATTCACATGAAAGAGACATCATGGGTA 160  
 QY 240 ACGGAAAAAATCATATGAGGATATTTTCAAAATACAGCCATATCAACGATATCCACT 299  
 Db 161 TAGAGAAATTTCCAT-----GAAAGCATCATTCACATCGAGATTTCCATTTATGG 214  
 QY 300 AAATATATCTCTGCGTATCCATTTCTTAAATGCTGTTAGTAACTACAGGACATGAT 359  
 Db 215 GACTATGAGCAAAATATCTATATGACAAATTGATATCCTTAGTAAATCATGCGGGCATGAT 274  
 QY 360 TAGAGAGATTT 370  
 Db 275 TATAGAGTTT 285

## RESULT 4

CB958289  
 LOCUS AGENCOURT\_13778124 NIH\_MGC\_184 Homo sapiens cDNA clone  
 DEFINITION IMAGE:30352112 5', mRNA sequence.  
 ACCESSION CB958289  
 VERSION CB958289.1 GI:30214405  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 792)  
 NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-r@mail.nih.gov  
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: http://image.llnl.gov

Plate: NDCM150 row: 1 column: 09  
 High quality sequence stop: 498.  
 FEATURES Location/Qualifiers  
 source 1. .792

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clones="IMAGE:30352112"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_184"  
 /notes="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site\_1: SfiI (ggccattggcc); Site 2: SfiI (ggccgctggcc); Library is oligo-dr primed and directionally cloned. cDNA was prepared from a glandular pool of tissues from thyroid, parathyroid, adrenal, cortex and pineal gland. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.38 kb (range 0.60-3.5 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 13.0%; Score 78.6; DB 14; Length 792;  
 Best Local Similarity 60.2%; Pred. No. 0.0014;  
 Matches 151; Conservative 0; Mismatches 94; Indels 6; Gaps 1;  
 QY 120 GAATTTTCATCTTTTCATGACTGGACTCCACCATAATATGAAGATCTTTATCTTTGTCTTCAT 179  
 Db 40 GACTCTCTCTTGGTAAAGGACTCAGCCAACTATGAAGTTTITGCTTTGCTTTAGT 99  
 QY 180 TATGGCTCTCATCCTAGCCTAGCTGATTTAGACTGATTCATCTGAAGAGAACTGACAGGAA 239  
 Db 100 CTTCGCTCTCATGATTTTCATGATTTAGCGCTGATTCACATGAAAGAGACATCATGGTA 159  
 QY 240 ACGGAAAAAATCATATGAGGATATTTTCAACATACAGCCATATCAAGATATCCACT 299  
 Db 160 TAGAGAAATTTCCATGAAAA-----CATCATTCATCGAGATTTCCATTTATGG 213  
 QY 300 AAATATATCTCTGCGTATCCATTTCTTAAATGCTGTTAGTAACTACAGGACATGAT 359  
 Db 214 GGACTATGGATCAAAATATCTATATGACAAATTGATATCCTTAGTAAATCATGCGGGCATGAT 273  
 QY 360 TAGAGAGATTT 370  
 Db 274 TATAGAGTTT 284

## RESULT 5

CB984475  
 LOCUS AGENCOURT\_13459298 NIH\_MGC\_184 Homo sapiens cDNA clone  
 DEFINITION IMAGE:30327061 5', mRNA sequence.  
 ACCESSION CB984475  
 VERSION CB984475.1 GI:30278999  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 793)  
 NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-r@mail.nih.gov  
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: http://image.llnl.gov

found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: NDCM135 row: h column: 14  
 High quality sequence stop: 486.  
 Location/Qualifiers

# FEATURES

1. 793  
 Location/Qualifiers

source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone="IMAGE:30327061"  
 /clone\_lib="NIH MGC 184"  
 /notes="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site 1: SfiI (ggcattatggcc); Site 2: SfiI (ggcgcctggcc); Library is oligo-dT primed and directionally cloned. cDNA was prepared from a glandular pool of tissues from thyroid, parathyroid, adrenal, cortex and pineal gland. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGGCGGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.38 Kb (range 0.60-3.5 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 13.0%; Score 78.6; DB 14; Length 793;  
 Best Local Similarity 60.2%; Pred. No. 0.0014;  
 Matches 151; Conservative 0; Mismatches 94; Indels 6; Gaps 1;  
 QY 120 GAATTCATCTTCATGACTGGACTCCACCAATATGAGATCTTATCTTCTTCAT 179  
 Db 41 GACTCTCTCTTGAGTAAAGGACTCAGCCAATGAGAGTTTTGTCTTGTCTTAGT 100  
 QY 180 TATGGCTCTCATCTGACCATGATTAGAGCTGATTCATCTGAAGAAACGTCACAGAA 239  
 Db 101 CTGGCTCTCATGATTTCCTATGATTAGCGTGATTCACATGAAGAAAGACATCATGGGTA 160  
 QY 240 ACGGAAACACATAGAGGATATTTCAACATACCGCCATACACGATATCCACT 299  
 Db 161 TAGAAGAAATCCAT-----GAAAGCATCATTCACATCCAGATTTCCATTTATGG 214  
 QY 300 AAATTATCTCTCGGTATCCATTTCTTAAATGCTGCTTAGTAATACTACAGGACATGAT 359  
 Db 215 GGACTATGGATCAATATCTATATGACAAATGATATCTTAGTATATCATGGGATGAT 274  
 QY 360 TAGAGAGATTT 370  
 Db 275 TATAGAGTTT 285

RESULT 6  
 CB985521  
 LOCUS  
 DEFINITION  
 AGENCOURT\_13652179 NIH\_MGC\_184 Homo sapiens cDNA clone  
 IMAGE:30327061 5', mRNA sequence.  
 ACCESSION  
 CB985521  
 VERSION  
 CB985521.1 GI:30280045  
 EST.  
 KEYWORDS  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens

REFERENCE  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 831)  
 NIH-MGC <http://imgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 JOURNAL  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
 CDNA Library Preparation: CLONTECH Laboratories, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: NDCM136 row: m column: 02  
 High quality sequence stop: 177.  
 Location/Qualifiers

# FEATURES

1. 831  
 Location/Qualifiers

source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30327553"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH MGC 184"  
 /notes="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site 1: SfiI (ggcattatggcc); Site 2: SfiI (ggcgcctggcc); Library is oligo-dT primed and directionally cloned. cDNA was prepared from a glandular pool of tissues from thyroid, parathyroid, adrenal, cortex and pineal gland. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGGCGGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.38 Kb (range 0.60-3.5 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 13.0%; Score 78.6; DB 14; Length 831;  
 Best Local Similarity 60.2%; Pred. No. 0.0013;  
 Matches 151; Conservative 0; Mismatches 94; Indels 6; Gaps 1;  
 QY 120 GAATTCATCTTCATGACTGGACTCCACCAATATGAGATCTTATCTTGTCTTCAT 179  
 Db 41 GACTCTCTCTTGAGTAAAGGACTCAGCCAATGAGAGTTTTGTCTTGTCTTAGT 100  
 QY 180 TATGGCTCTCATCTGACCATGATTAGAGCTGATTCATCTGAAGAAACGTCACAGAA 239  
 Db 101 CTGGCTCTCATGATTTCCTATGATTAGCGTGATTCACATGAAGAAAGACATCATGGTA 160  
 QY 240 ACGGAAACACATAGAGGATATTTCAACATACCGCCATACACGATATCCACT 299  
 Db 161 TAGAAGAAATCCAT-----GAAAGCATCATTCACATCCAGATTTCCATTTATGG 214  
 QY 300 AAATTATCTCTCGGTATCCATTTCTTAAATGCTGCTTAGTAATACTACAGGACATGAT 359  
 Db 215 GGACTATGGATCAATATCTATATGACAAATGATATCTTAGTATATCATGGGATGAT 274  
 QY 360 TAGAGAGATTT 370  
 Db 275 TATAGAGTTT 285

RESULT 7  
 CB987745  
 LOCUS  
 DEFINITION  
 AGENCOURT\_13631497 NIH\_MGC\_184 Homo sapiens cDNA clone  
 IMAGE:30327757 5', mRNA sequence.  
 ACCESSION  
 CB987745  
 VERSION  
 CB987745.1 GI:30282265  
 EST.  
 KEYWORDS  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens

REFERENCE  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 929)  
 NIH-MGC <http://imgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 JOURNAL  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits

CDNA Library Preparation: CLONTECH Laboratories, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: NDCM137 row: e column: 14  
 High quality sequence stop: 386.

## FEATURES

Location/Qualifiers  
 1. .929

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30327757"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH MGC 184"  
 /note="Organ: Pooled glandular; Vector: pDNR-LIB; Site\_1:  
 SfiI (ggccattggcc); Site\_2: SfiI (ggccctggcc);  
 Library is oligo-dT primed and directionally cloned. cDNA  
 was prepared from a glandular pool of tissues from thyroid,  
 parathyroid, adrenal, cortex and pineal gland. 5' and 3'  
 adaptors were used in cloning as follows: 5' adaptor  
 sequence: 5'-CACGGCATTATGGC-3' and 3' adaptor sequence:  
 5'-ATTCTAGAGCGGCGGCGGCGCATG-dT(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size 1.38  
 kb (range 0.60-3.5 kb). 15/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 13.0%; Score 78.6; DB 14; Length 929;  
 Best Local Similarity 60.2%; Pred. No. 0.0013;  
 Matches 151; Conservative 0; Mismatches 94; Indels 6; Gaps 1;

QY 120 GAATTTTCATCTTTTCATGACTGGACTCCACCAATATGAAGATCTTTATCTTTGCTTCAT 179  
 Db 21 GACTCTCTCTTGGTAAAGAGCTCAGCCAACTATGAAGTTTTTGTCTTTGCTTTAGT 80

QY 180 TATGGCTCTATCCTAGCAGTATGAGCTGATTCATCTGAAGAGAAAGCTCAGAGAA 239  
 Db 81 CTTGGCTCTCATCTTCATTTCCATGATTAGCGCTGATTCACATGAAAGAGACATCATGGGTA 140

QY 240 ACGGAAACATCATAGAGGATATTTTCAACAATACAGCCATATCAACGATATCCACT 299  
 Db 141 TAGAGAAATTCAT-----GAAAGCATCATTCACATCGAGATTTCCATTTATGG 194

QY 300 AAATTATCCCTCGGTATCCATTTCCCTTAAATGCTTGTAGTAACTACAGGACATGAT 359  
 Db 195 GGACTATGGATCAAAATATCTATATGACAAATGATATCCTTAGTAAATCATGGGCGATGAT 254

QY 360 TAGAGAGATTT 370  
 Db 255 TATAGAGGTTT 265

RESULT 8  
 BX485825 583 bp mRNA linear EST 04-SEP-2003  
 LOCUS  
 DEFINITION DKFZp686L09248\_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone  
 DKFZp686L09248 5', mRNA sequence.  
 ACCESSION BX485825  
 VERSION BX485825.1 GI:31948944  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 583)  
 Othenwaelder, B., Obermaier, B., Deutschenbaur, S., Mewes, H.W.,  
 Weill, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.  
 EST (Othenwaelder, B., Obermaier, B., Deutschenbaur, S., Mewes, H.W.,  
 et al.)

## JOURNAL COMMENT

Unpublished (2003)  
 Contact: MIPS  
 MIPS  
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
 This is the 5' sequence of the clone insert  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by MediGenomix (Martinsried/Germany) within the cDNA  
 sequencing consortium of the German Genome Project. No s1 sequence  
 available.  
 This clone (DKFZp686L09248) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.

## FEATURES

Location/Qualifiers  
 1. .583  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DKFZp686L09248"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="686 (synonym: hlcc3)"  
 /note="Vector: pTriplex2; Site\_1: SfiIA; Site\_2: SfiIB;  
 cDNA-collection"

## ORIGIN

Query Match 12.8%; Score 77.6; DB 13; Length 583;  
 Best Local Similarity 56.2%; Pred. No. 0.0023;  
 Matches 189; Conservative 0; Mismatches 139; Indels 8; Gaps 2;

QY 120 GAATTTTCATCTTTTCATGACTGGACTCCACCAATATGAAGATCTTTATCTTTGCTTCAT 179  
 Db 38 GACTCTCTCTTGGTAAAGAGCTCAGCCAACTATGAAGTTTTTGTCTTTGCTTTAGT 97

QY 180 TATGGCTCTATCCTAGCAGTATGAGCTGATTCATCTGAAGAGAAAGCTCAGAGAA 239  
 Db 98 CTTGGCTCTCATCTTCATTTCCATGATTAGCGCTGATTCACATGAAAGAGACATCATGGGTA 157

QY 240 ACGGAAACATCATAGAGGATATTTTCAACAATACAGCCATATCAACGATATCCACT 299  
 Db 158 TAGAGAAATTCAT-----GAAAGCATCATTCACATCGAGATTTCCATTTATGG 211

QY 300 AAATTATCCCTCGGTATCCATTTCCCTTAAATGCTTGTAGTAACTACAGGACATGAT 359  
 Db 212 GGACTATGGATCAAAATATCTATATGACAAATGATATCTTATGATCATGGGCGATGAT 271

QY 360 TAGAGAGATTTTTCACAAATGATTTTCTCTACTCTTTCTGTGTGTTGTTGAAAACCATCTTTC 419  
 Db 272 TATAGAG--GTAAGCTGACTCTAGTTGCTTTCTTTCTAGAAGTGTCAACACTGACAGTT 329

QY 420 AAATGAATATAAAACAAAGAAAAAAATCAAGTCAAGT 455  
 Db 330 TAAAAAAAAGCCATAGCTACACCATTCAGT 365

RESULT 9  
 BX485714 622 bp mRNA linear EST 04-SEP-2003  
 LOCUS  
 DEFINITION DKFZp686E20248\_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone  
 DKFZp686E20248 5', mRNA sequence.  
 ACCESSION BX485714  
 VERSION BX485714.1 GI:31948727  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 622)  
 Othenwaelder, B., Obermaier, B., Deutschenbaur, S., Mewes, H.W.,  
 Weill, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.  
 EST (Othenwaelder, B., Obermaier, B., Deutschenbaur, S., Mewes, H.W.,  
 et al.)  
 JOURNAL Unpublished (2003)

## COMMENT

Contact: MIPS  
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
 This is the 5' sequence of the clone insert  
 Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
 sequenced by MediGenomix (Martinsried/Germany) within the cDNA  
 sequencing consortium of the German Genome Project. No sl sequence  
 available.  
 This clone (DKFZp686E20248) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clones@rzpd.de.  
 Location/Qualifiers

## FEATURES

source

1. 622  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DKFZp686E20248"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="686 (synonym: hlcc3)"  
 /note="Vector: pTriplex2; Site\_1: SfiIA; Site\_2: SfiIB;  
 cDNA-collection"

## ORIGIN

Query Match 12.8%; Score 77.6; DB 13; Length 622;  
 Best Local Similarity 56.2%; Pred. No. 0.0023;  
 Matches 189; Conservative 0; Mismatches 139; Indels 8; Gaps 2;  
 QY 120 GAATTTTCATCTTTCATGACTGCTCCACCAATATGAGATCTTTATCTTTGCTTCAT 179  
 Db 36 GACTCTCTCTTGAGTAAAGGACTCAGCACTATGAGATTTTGTCTTTGCTTTAGT 95  
 QY 180 TATGCTCTCATCTAGCCATGATGAGTGTGATTCATCTGAGAGAAACGTCACAGGAA 239  
 Db 96 CTGGCTCTCATGATTTCCATGATTTAGCGTGTATTCACATGAAAGAGACATCGGTA 155  
 QY 240 ACGGAAACATCATAGAGGATATTTTCAACATACCGCCATATCAAGGATATCCACT 299  
 Db 156 TAGAGAAATTCAT-----GAAAGCATCATTCACATCGAGATTTCCATTTATGG 209  
 QY 300 AAATATCTCTCTCGGTATCCATTCCTTAAAGTGTGCTTGTAGTAACTACAGGACATGAT 359  
 Db 210 GGACTATGGATCAAAATATCTATATGACAATTCATATCTTCTAGTATCATCGGGCATGAT 269  
 QY 360 TAGAGAGATTTTTCACATGATTTTCTCTACTCTTTCTGTGTGAGAAACCATCTTTC 419  
 Db 270 TATAGAG--GTAAGCTGACTCTAGTACTTGTCTTTCTAGAGTGTCAACATGACGTT 327  
 QY 420 AAATGAATAAACAAGAGAAAAAATTCAGTCAAGT 455  
 Db 328 TAAAAAAGGCAATGCTAACAACCATTCAGT 363

## RESULT 10

CB987175  
 LOCUS CB987175 773 bp mRNA linear EST 01-MAY-2003  
 DEFINITION IMAGE:30329891 NIH\_MGC\_184 Homo sapiens cDNA clone  
 ACCESSION CB987175  
 VERSION CB987175.1 GI:30281695  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 773)  
 NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 AUTHORS Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgapbs@mail.nih.gov  
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits

cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: NDCMI42 row: n column: 12  
 High quality sequence stop: 181.  
 Location/Qualifiers

## FEATURES

source

1. 773  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30329891"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_184"  
 /note="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site\_1:  
 SfiII (ggccattggcc); Site\_2: SfiII (ggccctcgccc);  
 Library is oligo-dT primed and directionally cloned. cDNA  
 was prepared from a glandular pool of tissues from thyroid,  
 parathyroid, adrenal, cortex and pineal gland. 5' and 3'  
 adaptors were used in cloning as follows: 5' adaptor  
 sequence: 5'-CAGCGCCATATGGC-3' and 3' adaptor sequence:  
 5'-ATTCTAGAGCGGCGGCCGCAATG-dt(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size 1.38  
 kb (range 0.60-3.5 kb). 15/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 12.8%; Score 77.6; DB 14; Length 773;  
 Best Local Similarity 59.8%; Pred. No. 0.002;  
 Matches 150; Conservative 0; Mismatches 95; Indels 6; Gaps 1;  
 QY 120 GAATTTTCATCTTTCATGACTGCTCCACCAATATGAGATCTTTATCTTTGCTTCAT 179  
 Db 41 GACTCTCTCTTGAGTAAAGGACTCAGCACTATGAGATTTTGTCTTTGCTTTAGT 100  
 QY 180 TATGCTCTCATCTAGCCATGATTTAGAGTGTGATTCATCTGAGAGAAACGTCACAGGAA 239  
 Db 101 CTGGCTCTCATGATTTCCATGATTTAGCGTGTATTCACATGAAAGAGACATCGGTA 160  
 QY 240 ACGGAAACATCATAGAGGATATTTTCAACATACCGCCATATCAACGATATCCACT 299  
 Db 161 TAGAGAAATTCAT-----GAAAGCATCATTCACATCGAGATTTCCATTTATGG 214  
 QY 300 AAATATCTCTCTCGGTATCCATTCCTTAAAGTGTGCTTGTAGTAACTACAGGACATGAT 359  
 Db 215 GGACTATGGATCAAAATATCTATATGACAATTCATCTTAGTATCATCGGGCATGAT 274  
 QY 360 TAGAGAGATTT 370  
 Db 275 TATAGAGTTT 285

## RESULT 11

AA376700  
 LOCUS AA376700 312 bp mRNA linear EST 21-APR-1997  
 DEFINITION EST89157 Salivary gland Homo sapiens cDNA 5' end similar to  
 histatin 1, mRNA sequence.  
 ACCESSION AA376700  
 VERSION AA376700.1 GI:2029018  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 312)  
 Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,  
 Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D.,  
 White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C.,  
 Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Rhughes, J., Fine, L.D.,



Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

96026280

7566098

Other ESTs: EST89156 THC102726

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/cdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

## FEATURES

source

1..312

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="ATCC (inhost):181117"

/db\_xref="taxon:9606"

/dev\_stage="adult"

/clone\_lib="Salivary gland"

/note="Organ: salivary gland; Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI"

## ORIGIN

Query Match 12.7%; Score 77; DB 9; Length 312;

Best Local Similarity 59.8%; Pred. No. 0.004;

Matches 150; Conservative 0; Mismatches 95; Indels 6; Gaps 1;

QY 120 GAATTTTCATCTTCATGACCTGACCTCCACCAATATGAAGATCTTTATCTTGTCTTCAT 179

Db 12 GACTCTCTCTTCAGTAAAGACTCAGCCAACTATGAAGTTTGTCTTGTCTTAGT 71

QY 180 TATGGCTCTCATCTTAGCCATGATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGGAA 239

Db 72 CTTGGCTCTCATGATTTCCATGATTAGCGCTGATTCACATGAAAGAGACATCATGGGTA 131

QY 240 ACGGAAAAACATCATAGAGGATATTTCAACATACAGCCATATCAACGATATCCACT 299

Db 132 TAGAAGAAATTCAT-----GAAAAGCATCATTCACATCAGAGATTTCCATTTATGG 185

QY 300 AAATATTCCTCTCGCTATCCATTTCTTTAAATGCTGTGTAGTAACATACAGGACATGAT 359

Db 186 GGACTATGATCAAAATTAATCTATATGACAATTTGATATCCTTAGTAATCATGGGGCATGAT 245

QY 360 TAGAGAGATTT 370

Db 246 TATAGAGGTTT 256

## RESULT 12

AA382524

LOCUS

DEFINITION

EST95743 Testis I Homo sapiens cDNA 5' end similar to histatin 1,

mRNA sequence.

AA382524

ACCESSION

## VERSION

AA382524.1 GI:2034842

## KEYWORDS

Homo sapiens (human)

## SOURCE

Homo sapiens

## ORGANISM

Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 327)

## AUTHORS

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,K.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

96026280

7566098

Other ESTs: THC102726

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/cdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

## TITLE

Initial assessment of human gene diversity and expression patterns

## JOURNAL

Nature 377 (6547 Suppl), 3-174 (1995)

## MEDLINE

96026280

## PUBMED

7566098

## COMMENT

Other ESTs: THC102726

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/cdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

1..327

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="ATCC (inhost):186803"

/db\_xref="taxon:9606"

/sex="male"

/dev\_stage="adult"

/clone\_lib="testis I"

/note="Organ: testis; Vector: pBluescript SK-; Site\_1:

EcoRI; Site\_2: XhoI"

Query Match 12.7%; Score 77; DB 9; Length 327;

Best Local Similarity 59.8%; Pred. No. 0.0039;

Matches 150; Conservative 0; Mismatches 95; Indels 6; Gaps 1;

QY 120 GAATTTTCATCTTCATGACCTGACCTCCACCAATATGAAGATCTTTATCTTGTCTTCAT 179

Db 24 GACTCTCTCTTCAGTAAAGAGACTCAGCCAACTATGAAGTTTGTCTTGTCTTAGT 83

QY 180 TATGGCTCTCATCTTAGCCATGATTAGAGCTGATTCATCTGAAGAGAAACGTCACAGGAA 239

Db 84 CTTGGCTCTCATGATTTCCATGATTAGCGCTGATTCACATGAAAGAGACATCATGGGTA 143

QY 240 ACGGAAAAACATCATAGAGGATATTTCAACATACAGCCATATCAACGATATCCACT 299

Db 144 TAGAAGAAATTCAT-----GAAAAGCATCATTCACATCAGAGATTTCCATTTATGG 197

QY 300 AAATATTCCTCTCGCTATCCATTTCTTTAAATGCTGTGTAGTAACATACAGGACATGAT 359

Db 198 GGACTATGGATCAAAATTAATCTATATGACAATTTGATATCCTTAGTAATCATGGGGCATGAT 257



```

ORIGIN
  Site_1: EcoRI; Site_2: XhoI"
  Query Match      12.7%; Score 77; DB 9; Length 348;
  Best Local Similarity 59.8%; Pred. No. 0.0038;
  Matches 150; Conservative 0; Mismatches 95; Indels 6; Gaps 1;

QY 120 GAATTTTCATCTTTTCATGACTGGACTCCACCAATATGAAGATCTTTTATCTTTGTCTTCAT 179
    |||
DB 12 GACTCTCCCTTTGAGTAAAGGACTCAGCCACTATGAAGTCTTTTGTCTTTGTCTTTAGT 71
    |||

QY 180 TATGGCTCTCATCTAGCCATGATTAGAGTCAATCTGTAAGAGAAACGTCAACAGGAA 239
    |||
DB 72 CTTGGCTCTCATGATTTCATGATTAGCGTCAATTCACATGAAAGAGACATCATGGGTA 131
    |||

QY 240 ACGGAAAAACATCATAGAGGATATTTTCAACAATACACGACCATATCAACGATATCCACT 299
    |||
DB 132 TAGAAGAAATTCAT-----GAAAGCATCATTCACATCCAGAAATTTCCATTTATGG 185
    |||

QY 300 AAATTATCTCTCGCTATCCATTTCTTTAAATGCTGCTTAGTAACTACAGGACATGAT 359
    |||
DB 186 GGACTATGGATCAAAATATCTATATGACAATTGATATCCTTAGTAATCATGGGGCATGAT 245
    |||

QY 360 TAGAGAGATTT 370
    |||
DB 246 TATAGAGGTTT 256
    |||

RESULT 15
BX283872      390 bp      mRNA      linear      EST 05-MAR-2003
BX283872 NIH_MGC_81 Homo sapiens cDNA clone IMAGE958N191131 ;
IMAGE:4289874, mRNA sequence.
ACCESSION
EX283872
BX283872.1 GI:28848326
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 390)
Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Radelof, U., Schneider, D. and Korn, B.
Human UnigeneSet - RZPD3
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGE958N191131.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human UnigeneSet - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/ClonCards/cgi-
bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
pCMV-M13u, Primer sequence: CGTTGTAACACGAGCCAGT.
Location/Qualifiers
1..390
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE958N191131 ; IMAGE:4289874"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_81"
/note="Organ: muscle (skeletal); Vector: pDNR-LIB
(Clontech); Site: 1: SfiI (ggccatcgcc); Site: 2: SfiI
(ggcatatggcc); 5' and 3' adaptors were used in cloning
as follows: 5'-CACGCCCAATTATGGCC-3'

```

```

and 3' adaptor sequence:
5'-ATTCTAGAGCGCGGCGGCACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."

ORIGIN
  Query Match      12.7%; Score 77; DB 13; Length 390;
  Best Local Similarity 59.8%; Pred. No. 0.0036;
  Matches 150; Conservative 0; Mismatches 95; Indels 6; Gaps 1;

QY 120 GAATTTTCATCTTTTCATGACTGGACTCCACCAATATGAAGATCTTTTATCTTTGTCTTCAT 179
    |||
DB 35 GACTCTCTCTTTGAGTAAAGGACTCAGCCACTATGAAGTCTTTTGTCTTTGTCTTTAGT 94
    |||

QY 180 TATGGCTCTCATCTAGCCATGATTAGAGTCAATCTGTAAGAGAAACGTCAACAGGAA 239
    |||
DB 95 CTTGGCTCTCATGATTTCATGATTAGCGTCAATTCACATGAAAGAGACATCATGGGTA 154
    |||

QY 240 ACGGAAAAACATCATAGAGGATATTTTCAACAATACACGACCATATCAACGATATCCACT 299
    |||
DB 155 TAGAAGAAATTCAT-----GAAAGCATCATTCACATCGAGAAATTTCCATTTATGG 208
    |||

QY 300 AAATTATCTCTCTCGCTATCCATTTCTTTAAATGCTGCTTAGTAACTACAGGACATGAT 359
    |||
DB 209 GGACTATGGATCAAAATATCTATATGACAATTGATATCCTTAGTAATCATGGGGCATGAT 268
    |||

QY 360 TAGAGAGATTT 370
    |||
DB 269 TATAGAGGTTT 279
    |||

Search completed: August 16, 2004, 00:15:22
Job time : 2275 secs

```

FEATURES  
source

This Page Blank (uspto)